

Tietomallit

- Tietomallilla (data model) tarkoitetaan tiedon rakenteen ja tiedolle suoritettavan käsittelyn määrittävää kehikkoa - käsitteistöä
- Tietoa voidaan tarkastella eri näkökulmista - eri **abstraktiotasoilla**

1

Näkökulmat tietoon

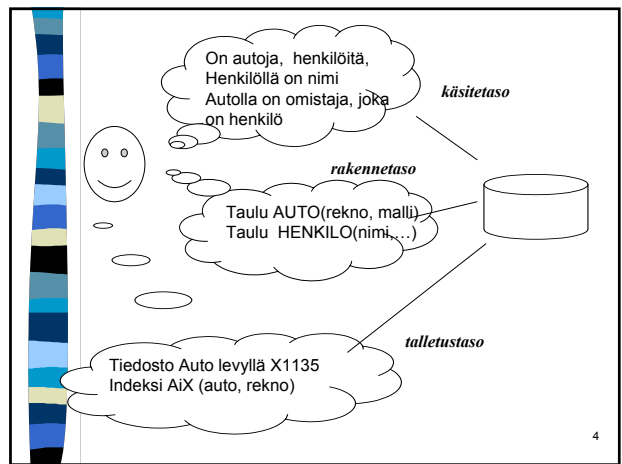
- Tiedon abstraktiotasot
 - **Käsitetaso**, kohdetaso (**conceptual level, real word level**):
 - Mitä tietoja käsitellään (millainen on se kohde, jota tiedoilla pitäisi kuvata; asiat, joita pitäisi esittää ja mitä tietoelementtien arvot tarkoittavat)
 - Miten tiedot liittyvät yhteen
 - Mitä sääntöjä tietoihin liittyy
 - Käsitellään JSS-kursseilla
 - **Rakennetaso**, looginen taso (**structural level, logical level**):
 - Minkälaisia käsiteltäviä rakenteita tiedot muodostavat - esim. eri ohjelmointikielten näkemykset tiedosta

2

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 ko kaikki keskitetyksi samassa paikassa.

3



4

Abstraktiotasot tiedon käsittelyssä

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

5

Mitä malleja olisi tarjolla?

käsitetaso	<ul style="list-style-type: none">• Entity-Relationship mallit (ER)• Oliomallit• Semanttiset tietomallit
rakennetaso	<ul style="list-style-type: none">• Relaatiomalli• Oliomallit• Hierarkkinen malli• Verkkomalli
fyysinen taso	<ul style="list-style-type: none">• Toimittajakohtaiset käsitteet

6

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.

7

Relaatiomalli

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

8

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

9

Relaatio vielä havainnollisemmin

Relaatiokaavion nimi

Sarakkeet

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

10

Relaation matemaattinen määritelmä

- Olkoon D_1, D_2, \dots, D_n arvojoukkoja, joiden ei tarvitse olla erillisiä.
- Relaatio R on joukko 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.

11

Monikko

- Monikko on 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).

12

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!

13

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

14

Relaatiokaavio

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

15

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 määrittystä käytettäessä oletetaan, että lukija osaa kuvaavien attribuuttinimien perusteella arvata niiden arvojoukot ja tulkinnat
- Esim. Auto(Reknro, Väri, Vuosimalli)

16

Relaatiokaavio

- Täydellisempi esitys:
Auto(Reknro: Suomalaiset_rekisterinumero, Väri: Autovärit, Vuosimalli: Vuosiluvut >1900)
- Relaatio on relaatiokaavion **ilmentymä**.
 - Relaatiokaavion ilmentymät kuvaavat jotain todellisuuden ilmiötä, vaikkapa autojen olemassaoloa tietyllä hetkellä, esim 1.1.2003
 - Toista ajankohtaa kuvaa eri relaatio, joskin saman kaavion ilmentymä sek in.

17

Relaatiokaavio

- Relaatiotietokannassa on tyypillisesti kustakin relaatiokaaviosta vain vain yksi ilmentymä kerrallaan
- Ilmentymään viitataan kaavion nimellä
- Tietokannan ylläpidossa kaavionimellä osoitettava ilmentymä vaihdetaan toiseksi

18

Relaatiokaavio

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

19

Relaatiokaavion käsitteitä

- Relaation koko (**cardinality**)
 - relaation monikoiden lukumäärä
- Relaation aste (**degree**)
 - relaatiokaavion attribuuttien lukumäärä
- Relaatiotietokanta (**relational database**)
 - sisältää yleensä useita relaatioita
- Relaatiotietokantakaavio (**relational database schema**)
 - relaatiotietokannan relaatiot määrittelevien relaatiokaavioiden kokoelma

20

Teoria vs havaintoesitys

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

21

Avain (key)

- Relaation monikot ovat keskenään erilaisia
 - Monikko pystytään identifioimaan sisältönsä perusteella
 - Identifiointiin ei välttämällä tarvita kaikkien attribuuttien arvoja
- Relaation **avain** (key) on sellainen attribuutti tai niiden yhdistelmä, jolle pätee
 - missään relaatiokaavion ilmentymässä ei voi olla kahta tai useampaa monikkoa, joissa on samat avain-attribuuttien arvot
 - yhdistelmästä ei voi poistaa yhtään attribuuttia siten että jäljellejäävät täyttäisivät edellisen ehdon (minimaalisuus)

22

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 ilmentymässä.
- Vmalli ei kuitenkaan erottele kaikissa mahdollisissa ilmentymissä.

23

Avain

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

24

Avain

- Pääavain esitetään relaatiokaaviossa alleviivaamalla siihen kuuluvat attribuutit

Auto (Rekno, Väri, Vuosimalli)

Työntekijä (Työntekijänumero,...)

Pelivaraus (kenttänumero, alkuaika, kesto, nimi)

25

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 monikkoihin
- Relaatiotietokannassa kytkentä saadaan aikaan sisällyttämällä monikkoon kytkettävän, siis vieraan, monikon (pää)avain.

26

Viiteavain

Auto	Rekno	Väri	Vmalli
	ISO-795	Musta	1992

Omistus	Henkilötunnus	Auto
	121212-1234	ISO-795

27

Viiteavain

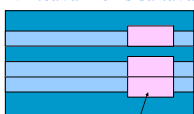
- Attribuuttia tai attribuuttiyhdistelmää, jonka arvot toteuttavat viittauksen toiseen monikkoon kutsutaan viiteavaimeksi (foreign key).
- Viiteavaimessa voidaan sallia myös tyhjäärvot – tällöin relaation monikkoja ei tarvitse kytkeä mihinkään

28

Viiteavain

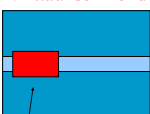
- Vaikka viiteavain viittaakin yhteen suuntaan sitä voi hyödyntää kaksisuuntaisesti esim. selvittäessä :
 - mikä auto kytkeytyy tiettyyn omistukseen
 - mitkä omistukset kytkeytyvät tiettyyn autoon

Viiteavaimen sisältävä



'kopiot'

Viittauksen kohde



'originaali'

29

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

30

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	TyöntekijäNumero	...	Esimies
1010			
1020			1010
1030			1010

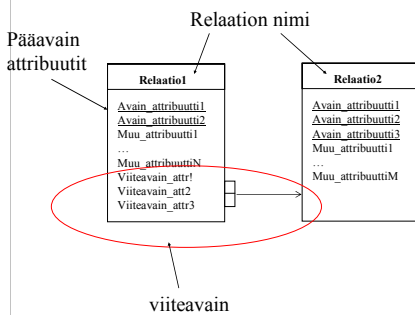
31

Viiteavain

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

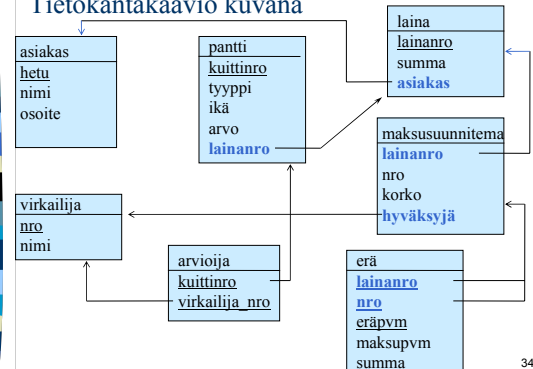
32

Tietokantakaavio kuvana



33

Tietokantakaavio kuvana



34