

P114

Klasické metody modelování

RDM, ERAM

2

Témata

- modelování v RDM
- univerzální relace
- dekompozice, normalizace
- syntéza relací
- omezení DM v RDM
- modelování v ERAM
- notace, postup
- špatně a správně utvořené modely

Cílem je implementace:

- Co máme na počítači?
 - soubory / tabulky
 - záznamy / řádky
 - položky / sloupce,
 - klíče
 - co se jak implementuje
(příklad tabulek Rozvrhu)

ROZVRH

UCITEL

#UC		

PREDMET

#PR		

MISTNOST

#MI		

#UC	#PR	#MI	CAS

- popisy objektů
- zachycení souvislostí
- použití klíčů

Modelování v RDM

- definice relace, schématu relace
- klíče, primární, alternativní, cizí, referenční integrita
- relační schéma (databáze)
- funkční závislosti, Armstrongova pravidla
- objekty a vztahy jako relace

OPAKOVÁNÍ:

definice relace, schématu relace

- R je subset $D_1 \times \dots \times D_n$
- $R(A_1:D_1, \dots, A_n:D_n)$, $D_i = \text{dom}(A_i)$
schéma relace
 $\mathbf{A} = \{A_1:D_1, \dots, A_n:D_n\}$ množina atributů relace R ,
 $R(\mathbf{A})$ jiný zápis schématu relace R nad množinou atributů \mathbf{A}
- schéma relace = záhlaví tabulky
n-tice = řádek tabulky
- rozdíly:
tabulka má vždy nějaké pořadí řádků a sloupců
tabulka může mít duplicitní řádky

OPAKOVÁNÍ: projekce

- projekce n -tice na podmnožinu atributů:
 B je subset A , u je n -tice z R : $u[B]$ je projekce
(k -tice s komponentami z B)
- projekce relace na podmnožinu atributů:
je to projekce všech n -tic z R na podmnožinu atributů B :
 $R[B]$

OPAKOVÁNÍ: klíče, primární, alternativní, cizí, referenční integrita

- klíč K relace $R(\mathbf{A})$: $K \subseteq \mathbf{A}$, u, v jsou z R různé: $u[K] \neq v[K]$, if K' je subset \mathbf{A} a má tutéž vlastnost jako K , pak K' obsahuje K
- kandidáti na prim. klíč, primární klíč - jeden zvolený, ostatní kandidáti: alternativní klíče
- jednoduché a složené klíče
- cizí klíč $C_K :=$ skupina atributů, která je primárním klíčem K jiné relace
- referenční integrita: $R_2[C_K]$ je subset $R_1[K]$

OPAKOVÁNÍ: relační schéma (databáze)

- **RSD** := **(R,I)**, kde **R** = {**R**₁, ... , **R**_m}, **I** je množina IO (logických podmínek, které musí data v DB splňovat)
- lokální IO: omezují data v jednom schématu relace
- globální IO: dávají vazby mezi daty různých schémat relací
- přípustná relační databáze R se schématem **(R,I)**
- stav databáze R

OPAKOVÁNÍ:

funkční závislosti, Armstrongova pravidla

- funkční závislost je vztah mezi daty v „tabulkách“
- funkční závislost je druhem IO
- B, C jsou subsety A : $B \rightarrow C$ jestliže pro libovolné n -tice $u, v \in R$ platí if $u[B] = v[B]$ then $u[C] = v[C]$
- X, Y, Z jsou subsety A . Potom:
 - if $Y \subseteq X$ then $X \rightarrow Y$ (triviální závislost) (AP1)
 - if $X \rightarrow Y$ and $Y \rightarrow Z$ then $X \rightarrow Z$ (AP2)
 - if $X \rightarrow Y$ and $X \rightarrow Z$ then $X \rightarrow YZ$ (AP3)
 - if $X \rightarrow YZ$ then $X \rightarrow Y$ and $X \rightarrow Z$ (AP4)

OPAKOVÁNÍ: normální formy

- 1NF: domény obsahují pouze atomické prvky (nikoli znovu relace)
- 2NF: 1NF + neexistují parciální fční závislosti na klíči
- 3NF: 2NF + neexistují transitivní funkční závislosti (C tranzitivně závisí na X: $X \rightarrow Y \rightarrow C$ a $C \notin X, C \notin Y$, a $Y \not\rightarrow X$)
- BCNF pro každou netriviální závislost $X \rightarrow Y$ platí X obsahuje klíč schématu relace R

OPAKOVÁNÍ:

dekompozice, normalizace, syntéza

- (pragmatické) důvody pro zavedení xNF: aktualizací anomálie
- normalizace pomocí dekompozice relačních schémat (použití AP4)
- konstrukce relačního schématu syntézou (použitím funkčních závislostí - AP3)

Univerzální relace

- modelování dekompozicí univerzální relace
- předpoklad schématu univerzální relace:
jednoznačnost jmen atributů
- předpoklad jednoznačnosti vztahů
 - protipříklad:
Ved_diplomky(Učitel, Student)
Učí (Učitel, Předmět, Student)
nelze získat z jednoho schématu univerzální relace

Omezení DM v RDM

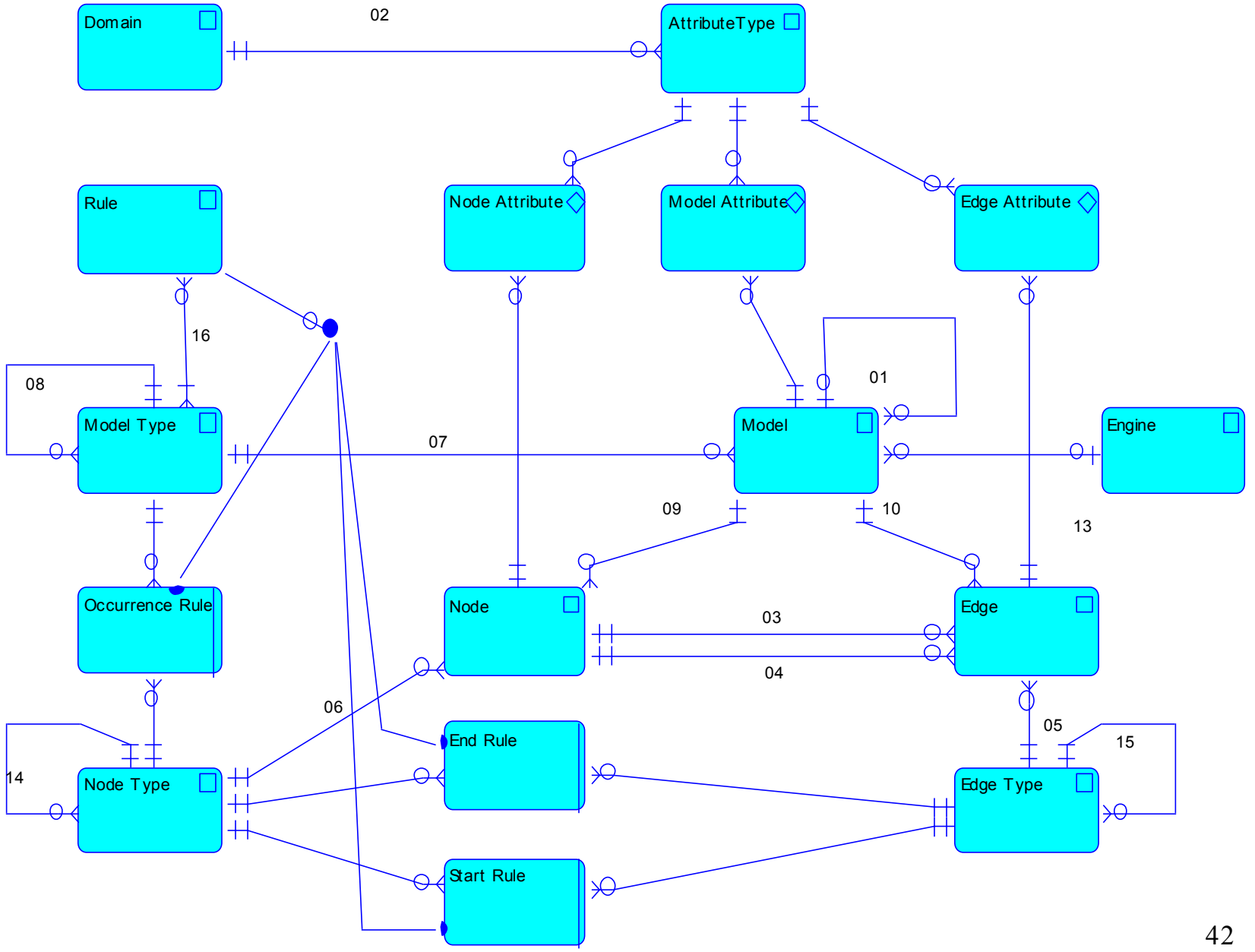
- praktická tvorba datového modelu v RDM
 - kombinace dekompozice a syntézy
- absence sémantiky ve formálním aparátu
 - hra symbolů jejichž interpretace leží „za hranicemi systému“
- nedostatečnost PL1 pro analýzu přirozeného jazyka
 - potřebujeme v jednom systému pracovat s objekty různých řádů
- přílišná formalizace snižující využitelnost intuice
 - o čem se vyjadřujeme, versus
 - o čem skutečně přemýšlíme
- nereálné předpoklady a jejich obcházení „krokem stranou“
 - předpoklad jednoznačnosti vztahů

Modelování v ERAM

- objekt -- kontejnerem je „entita“, „typ entity“
- vztah -- kontejnerem je „typ vztahu“
- atribut (typu entity nebo typu vztahu) -- je funkce přiřazující hodnoty popisných typů
- IO --soulad schématu s modelovanou realitou
- diagram typů entit a vztahů ERD
- kardinalita vztahu, členství ve vztahu
- ISA vztah

Notace (nástroje)

- entita
 - silná, popisná, vazební
 - klíče (primární, alternativní, cizí, nejednoznačné)
- vazba
 - maximální kardinalita, minimální kardinalita
 - role MASTER, DETAIL
 - pojmenování vazeb
- podtypy entity
 - definice podtypu, dědění
 - skupiny disjunktních podtypů s úplným nebo částečným pokrytím



Postup (kroky)

- identifikuj entity
- urči vazby (mřížka entit)
- vytvoř model: doplň diagram + kardinality
- odstraň duplicitní vazby
- modeluj podtypy entit a uprav vazby
 - totalita atributů, parcialita vazeb, rozdělení do kategorií
 - stav entity, rekurzivní vazby, vazba typu kusovník
- Příklad: Rozvrh

„Nesprávnost“ modelu, příklady

- diagram bez sémantiky
- vztahy jsou v algoritmech a ne ve vazbách
 - Pojišťovací systém (rozvoje-schopnost „-“)
 - SELECT SE (procesy) (důsledek pro slévání submodelů)
- absence nadtypu a tím velká složitost vazeb
 - Údržba v REAS (velká složitost vazeb)
 - Pojišťovací systém (nekonzistentnost informací o partnerech)
- konečný počet podtypů
 - a algoritmů jejich zpracování (problém přidání podtypu, změny klasifikace)

„Správný“ model, příklady

- co je to „správný model“
 - pozor na „reálný svět“ vs „svět představ“
 - adekvátnost požadavkům (i nevyjádřeným)
 - úloha analytika - „datového modeláře“
 - text může být blábol ze špatně utvořených vět
 - model může být špatně konstruován z nedobře definovaných konstruktů
- „správný“ model:
 - IS Bílý Motýl
 - transakční systém EXPANDIA Banky