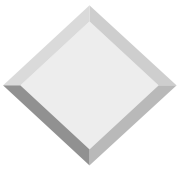


Konceptuální analýza a návrh

Jaroslav Pokorný
MFF UK, Praha
pokorny@ksi.ms.mff.cuni.cz

Obsah

- úvod
- metodologie návrhu
- kontrola kvality a úplnosti E-R schématu
  - pravidla pro kreslení E-R diagramů
  - seznamy kontrolních otázek
- strategie v analýze a návrhu
  - shora dolů
  - zdola nahoru
  - zevnitř ven
  - smíšená
  - integrace schémat

1. Úvod

Cíl: návrh konceptuálního a databázového schématu, návrh procesů.

Předpoklad: původní varianta E-R modelu (Chen, 1976), diagramy toků dat

Cílové modely: RMD resp. tabulky v SQL

Pomůcka: metodologie návrhu

1. Úvod

cílem *metodologie konceptuálního návrhu* je:

- upřesnit pravidla pro rozlišení mezi entitami, atributy a vztahy,
- upřesnit kardinality vztahů,
- dát pravidla pro vztahy, které mohou mít atributy,
- rozhodnout jak řešit problémy návrhu,
- podat způsob, jak popsat sémantiku aplikace prostřednictvím datových abstrakcí (konstruktů), které jsou k dispozici,
- dát pravidla pro volbu primárních klíčů z množiny kandidátů na klíč, včetně situace, kdy typ entity vyžaduje další identifikaci,

1. Úvod

- specifikovat detailní proceduru pro transformaci konceptuálního schématu v E-R modelu do RMD,
- specifikovat interpretaci prázdných hodnot cizích klíčů v entitní relaci.

Pz: ve skutečnosti jde o dvě schémata:

výsledek analýzy,

výsledek návrhu.

V praxi je obvyklé, že se mohou lišit.

1. Metodologie návrhu - jedna možnost

krok 1:

- definice typů nezávislých (referenčních) entit,
- definice typů vztahů,
- definice typů závislých entit, formulují se některé (většinou klíčové) atributy entit
- iterace předchozího.

krok 2:

- formulují se atributy,
- určují se identifikační klíče entit.

Návrh typů entit a atributů

Při definici atributů by se

(i) měly eliminovat vícehodnotové atributy,

(ii) vyjasnit rozdíly mezi typy entit a vztahů a atributy.

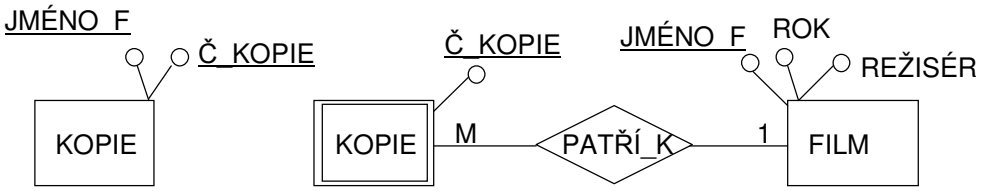


schéma a) schéma b)

Problém: identifikace slabé entity

Návrh typů entit a atributů

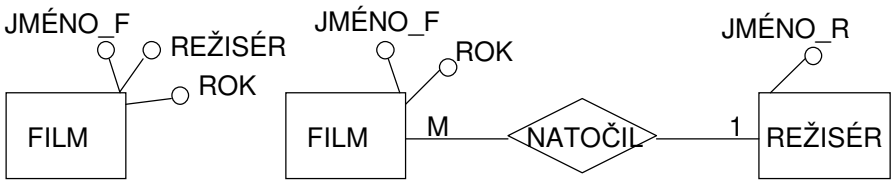


schéma a) schéma b)

Problém: nejde o atribut, ale o vztah

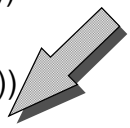
Návrh typů vztahů

Problém: úplnost E-R schématu

Situace v návrhu:
VEDE(REŽISÉR:(1,n), HEREC:(1,1)),
REŽÍRUJE(REŽISÉR:(1,n), FILM:(1,1)).
Zřejmě by měl existovat typ vztahu mezi HEREC a FILM.

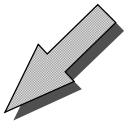
Situace v návrhu:
JE_NAPLÁNOVÁN_NA(HEREC:(0,n), FILM:(0,m)),
VYŽADUJE(SPECIÁLNÍ_ZNALOSTI:(0,n), FILM:(0,m)). asi schází

OVLÁDÁ(SPECIÁLNÍ_ZNALOSTI:(0,n), HEREC:(0,m))



Návrh typů vztahů

Situace v návrhu:
ZAMĚSTNANEC JE OSOBA,
REŽISÉR JE ZAMĚSTNANEC
REŽISÉR JE OSOBA



je nadbytečný

Problém: redundance v E-R schématu

Kontrola kvality E-R schématu

Pravidla pro kreslení ER diagramů

- preferovat horizontálně kreslené vztahy (je obvyklé číst je zleva doprava),
- u kardinalit 1:N preferovat typ entity s N nalevo (od častějších entit k méně častým),
- minimalizovat šikmé čáry a křížení čar (křížení lze vyjádřit úhlem nebo obloučkem),
- nekreslit paralelní čáry těsně u sebe (zvyšuje to přehlednost čtení diagramu),
- pro konkrétní řešení nějakého problému používat podmnožiny E-R diagramů,
- označit každý E-R diagram jménem, datem a identifikací autora.

Kontrola kvality E-R schématu

Metoda seznamů kontrolních otázek (check lists)

- TYPY ENTIT
  - je jméno typu entity smysluplné a je v jednotném čísle?
  - jak je to s výlučností?
  - jsou dány alespoň dva atributy?
  - není zvláštní, že existuje více než 8 atributů?
  - co homonyma a synonyma v názvech typů entit?
  - je definice typu entity úplná?
  - jsou známy nějaké skutečné četnosti entit daného typu?
  - je dán klíč entity?

Kontrola kvality E-R schématu

- existuje alespoň jeden typ vztahu, ve kterém je typ entity členem?
- existuje alespoň jeden proces, který vytváří, vybírá, aktualizuje, odstraňuje, archivuje a používá danou entitu?
- jsou dány požadavky na distribuci?
- mění se entita v čase?
- vyhovuje definice atributů požadavkům na normalizaci?
- není typ entity příliš generický?
- je typ entity naopak generický postačujícím způsobem?

Kontrola kvality E-R schématu

- PODTYPY ENTIT
- jsou podtypy typu entity výlučné?
- mají atributy a/nebo participují ve vztazích?
- mají vlastní identifikátor nebo přijímají identifikátor nadřazeného typu entity?
- je množina podtypů úplná?
- neměly by být podtypy modelovány alternativně?
- znáte atributy a/nebo vztahy a podmínky, které odlišují podtyp entity od jiného typu?

Kontrola kvality E-R schématu

- ATRIBUTY
  - má atribut jméno v jednotném čísle?
  - nezahrnuje jméno atributu jméno typu entity?
  - je atribut jednoduchý?
  - je dána definice formátu, délky, domény, odvození atd.?
  - nejde ve skutečnosti o chybějící typ entity?
  - nejde ve skutečnosti o chybějící typ vztahu?
  - není atribut replikován odjinud (např. z nadtypu)?
  - není důležité znát jeho hodnoty v čase?
  - závisí hodnota atributu pouze na entitě?

Kontrola kvality E-R schématu

- je-li hodnota atributu povinná, je vždy známá?
- potřebujeme pro atribut zavést doménu?
- je hodnota atributu funkčně závislá na části klíče?
- je hodnota atributu funkčně závislá na atributech, které nejsou částí klíče?

Kontrola kvality E-R schématu

- TYPY VZTAHŮ
  - je každý konec spojnice typu vztahu v E-R diagramu pojmenovaný?
  - má typ vztahu právě dva konce (spojnice)?
  - při použití invertované syntaxe, řekneme-li větu popisující vztah "inverzně", je vztah stále korektní?
  - má každý člen typu vztahu kardinalitu a typ členství?
  - je konstrukce platná?
  - nejde o zřídka se vyskytující konstrukce?
  - není typ vztahu redundantní?

Kontrola kvality E-R schématu

- TYPY VÝLUČNÝCH VZTAHŮ
  - vycházejí konce výlučných vztahů ze stejného typu entity?
  - má tento typ entity ve výlučných typech vztahů vždy stejný typ členství?
  - je vztah jen v jedné skupině výlučných vztahů?

Strategie v analýze a návrhu

Lze rozlišit 4 základní strategie:

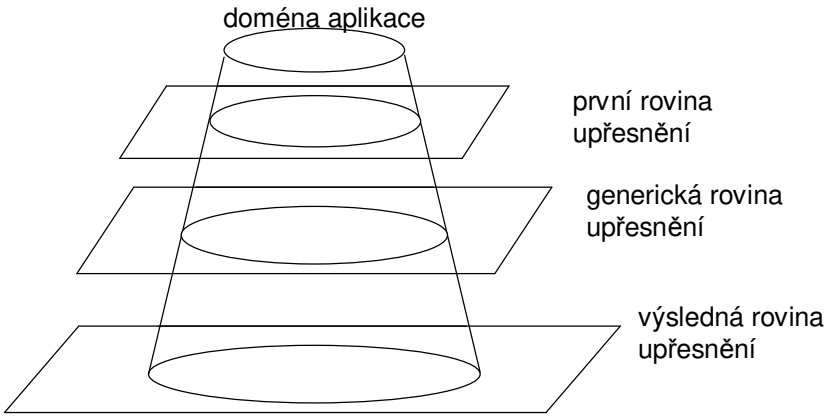
- shora dolů,
- zdola nahoru,
- zevnitř ven,
- smíšená.

Strategie návrhu

Příklad:

Databáze CENSUS s týká vzorku obyvatelstva. Osoby mají jméno a příjmení, pohlaví věk, místo narození, bydliště a délku pobytu v něm. U mužů se navíc sleduje stav (ženatý, …), u žen rodné jméno. Místa mohou být buď města (v tuzemsku) nebo státy (u cizinců). U míst se sleduje celková populace. Města se navíc přiřazují krajům.

Strategie návrhu shora dolů



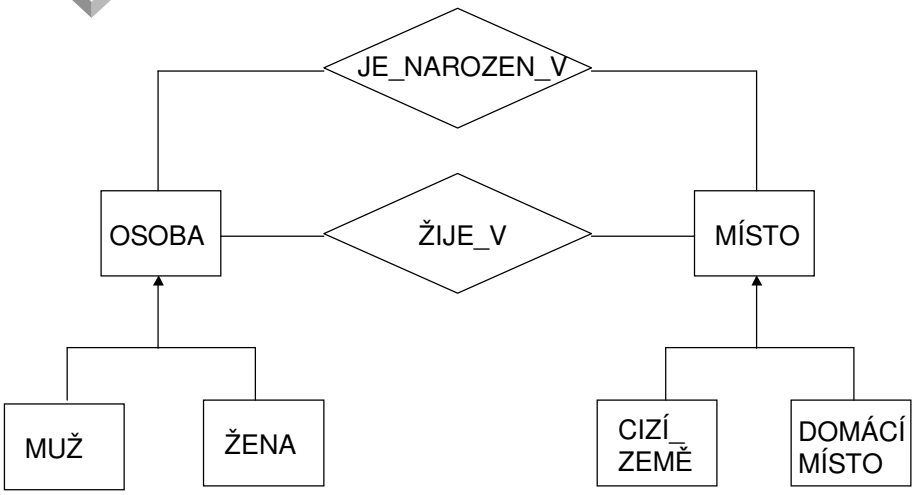
v praxi existuje více rovin upřesnění

Strategie návrhu shora dolů - 1. upřesnění

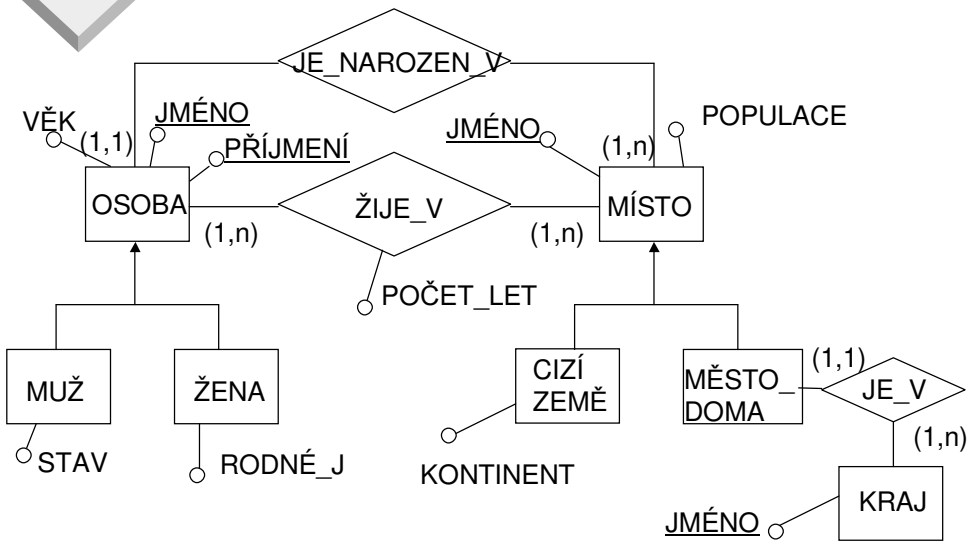
DEMOGRAFICKÁ _DATA

Strategie návrhu shora dolů - 2. upřesnění

Strategie návrhu shora dolů - 3. upřesnění



Strategie návrhu shora dolů - 4. upřesnění



Strategie návrhu shora dolů - zhodnocení

- mentálně dosti náročná
  - klade na projektanty vysoké nároky,
  - vyžaduje vysoký stupeň abstrakce,
  - schopnost odhadnout dobře globální požadavky
- složitá zvláště u rozsáhlých systémů.
- podporou může být i centralizované řízení organizace, pro kterou se IS vytváří.

Strategie návrhu zdola nahoru – množina výchozích atributů

JMÉNO_MUŽE

VĚK_ŽENY VĚK_MUŽE

JMÉNO_ŽENY

PŘÍJMENÍ_MUŽE

PŘÍJMENÍ_ ŽENY

STAV

RODNÉ_JMÉNO

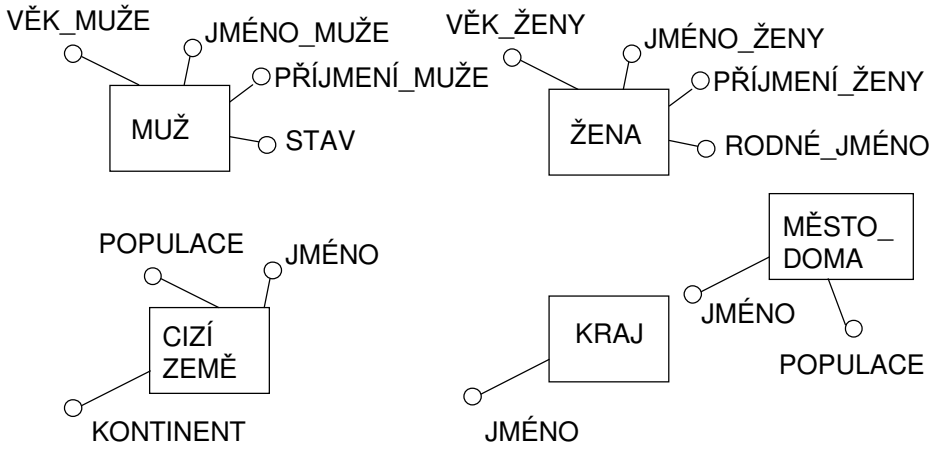
JMÉNO_CIZÍ_ZEMĚ

POPULACE_V_CIZÍ_ZEMI

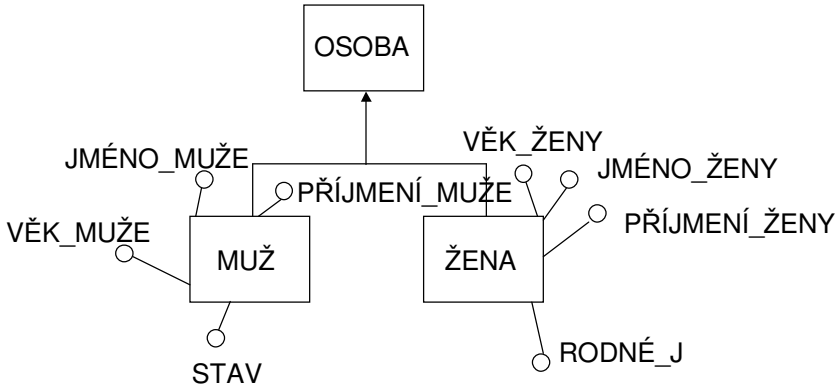
JMÉNO_MĚSTA_DOMA POPULACE_MĚSTA_DOMA

JMÉNO_STÁTU

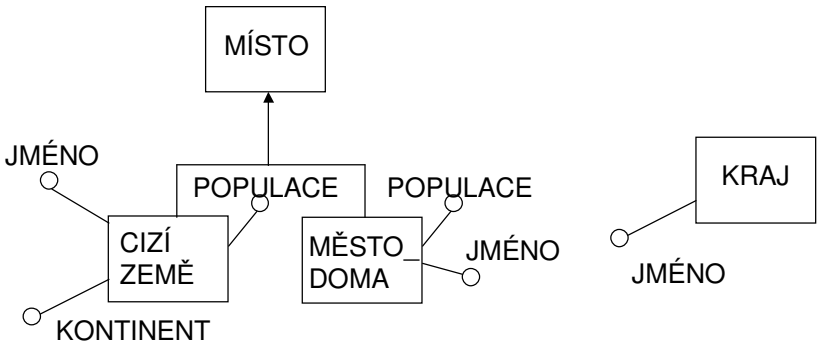
Strategie návrhu zdola nahoru – konstrukce typů entit



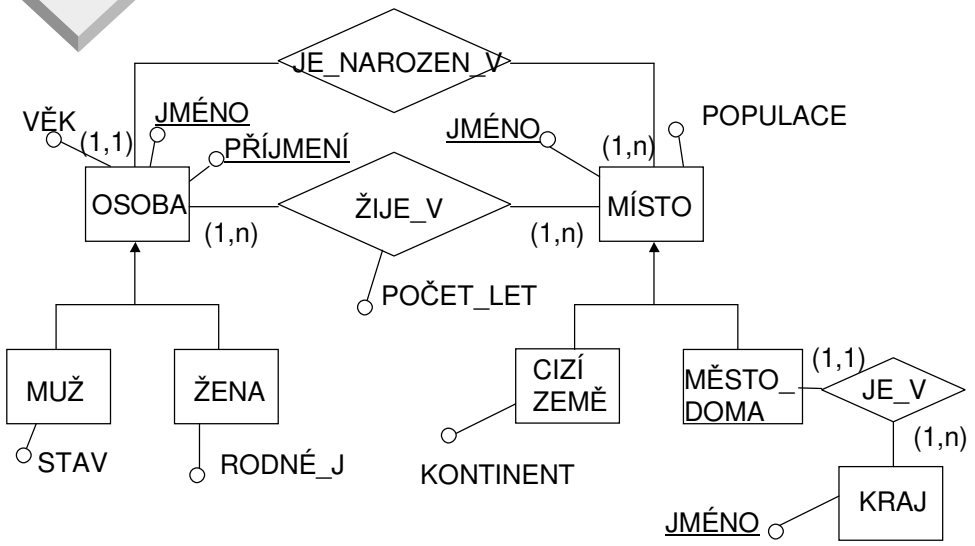
Strategie návrhu zdola nahoru –množina hierarchií typů entit



Strategie návrhu zdola nahoru –množina hierarchií typů entit



Strategie návrhu shora dolů - 4. upřesnění



Strategie návrhu zdola nahoru – zhodnocení

Výhody:

- jednoduchost,
- rychlost návrhu (možnost paralelního projektování).

Nevýhody:

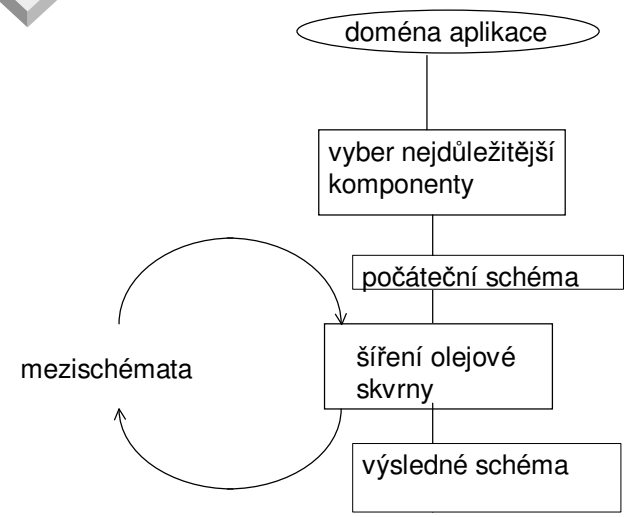
- častější restrukturalizace schémat.
- na počátku nepřehledné množství typů dat. Řada pojmů se objeví až na konci procedury návrhu.

metafora - les/stromy:

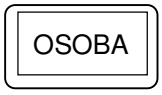
shora dolů: vidíme les a nevidíme stromy,

zdola nahoru: vidíme stromy a nevidíme les.

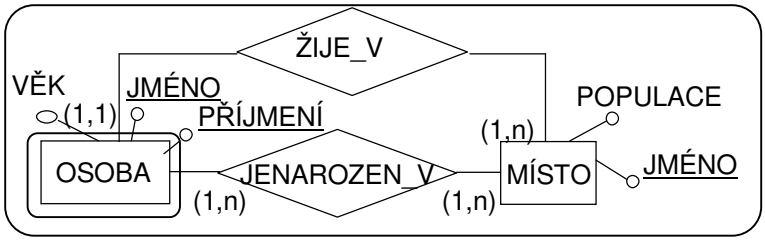
Strategie zevnitř ven



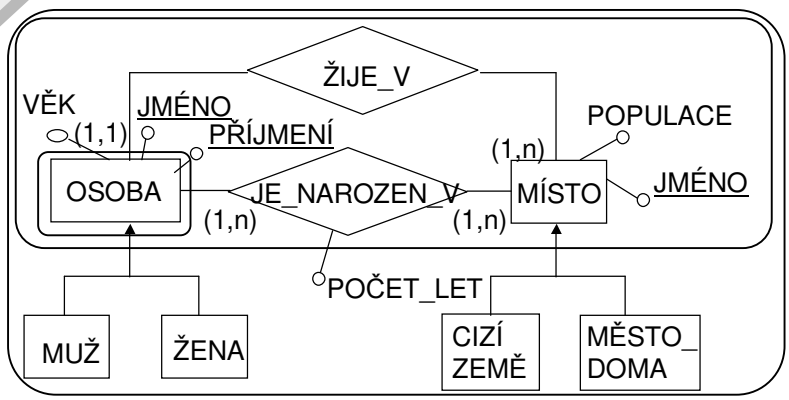
Strategie zevnitř ven - 1. stadium



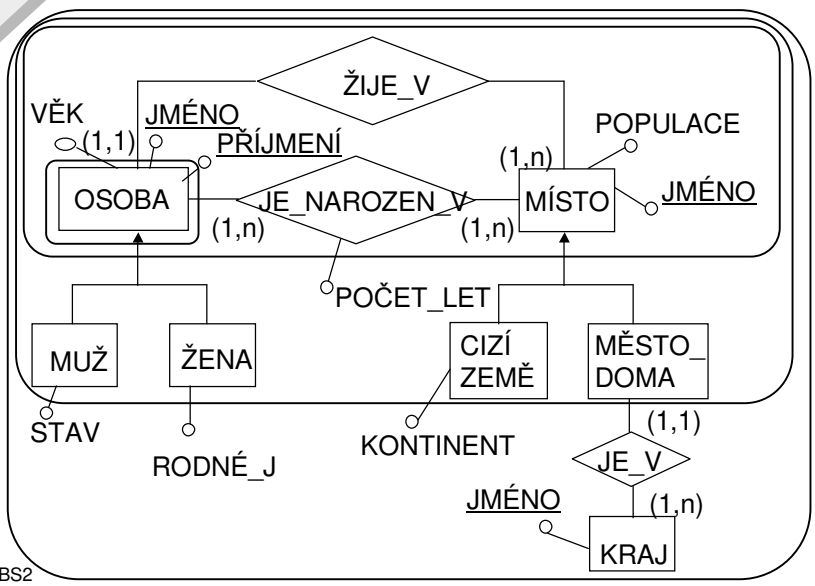
Strategie zevnitř ven - 2. stadium



Strategie zevnitř ven - 3. stadium



Strategie zevnitř ven - 4. stadium



Strategie zevnitř ven - zhodnocení

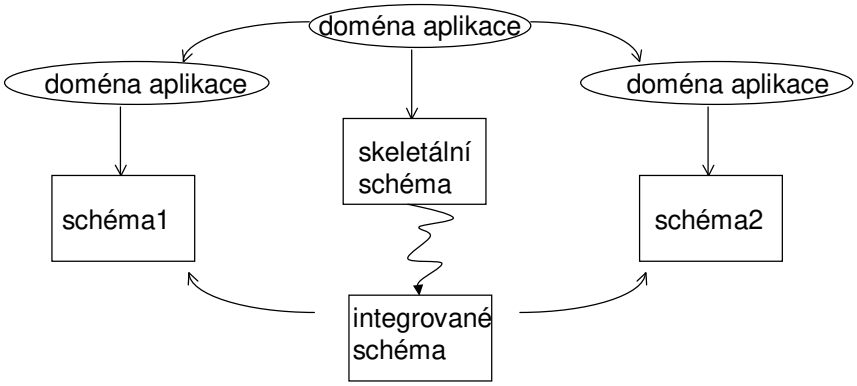
Výhody:

- jistá disciplína, podobná návrhu shora dolů.

Nevýhody:

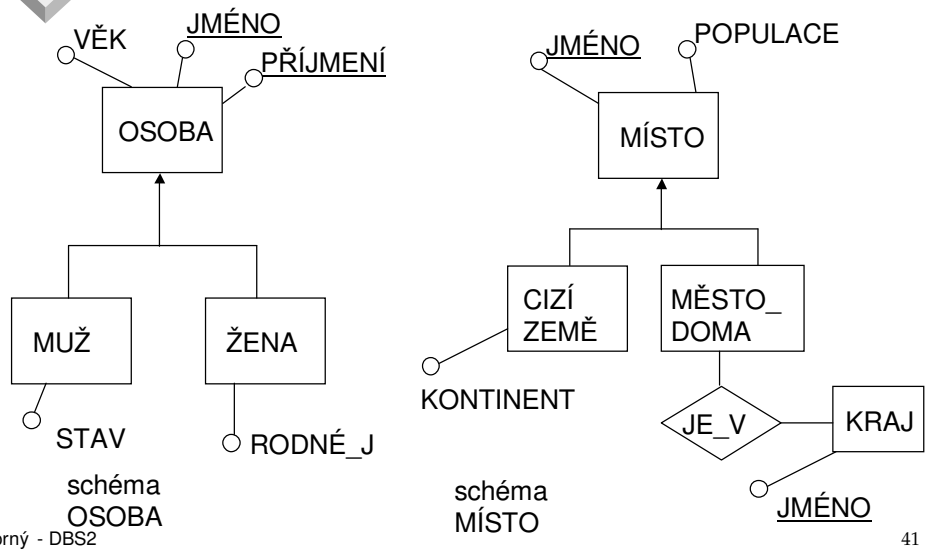
- oproti metodě shora dolů se pohybujeme pouze na jedné rovině abstrakce.

Smíšená strategie

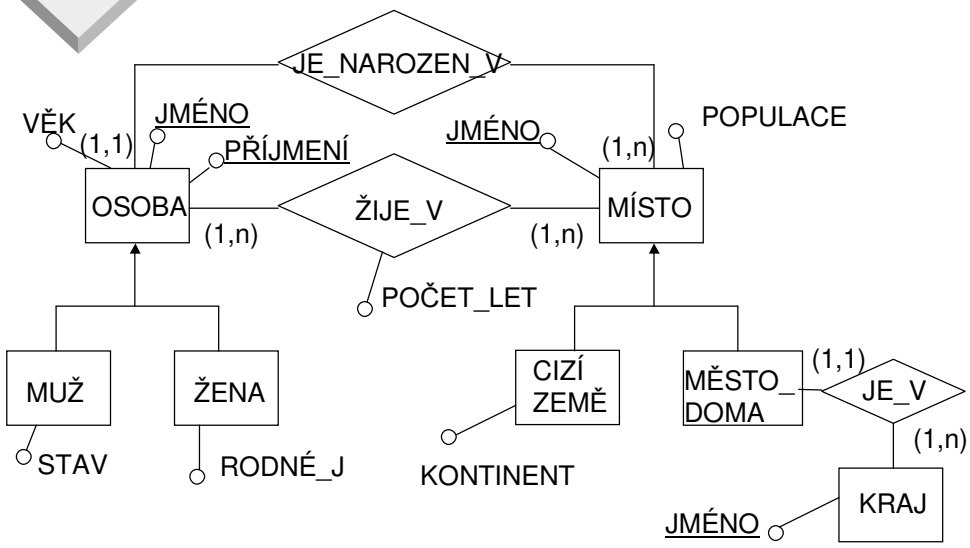


Smíšená strategie - skeletální schéma

Smíšená strategie - rozvoj dvou dílčích schémat



Strategie návrhu shora dolů - 4. upřesnění



Smíšená strategie - zhodnocení

Výhody:

- možnost souběžné práce
- použití různých strategií v jednotlivých schématech

Nevýhody:

- integrace schémat
- konflikty (*strukturální* a *ve jménech*)

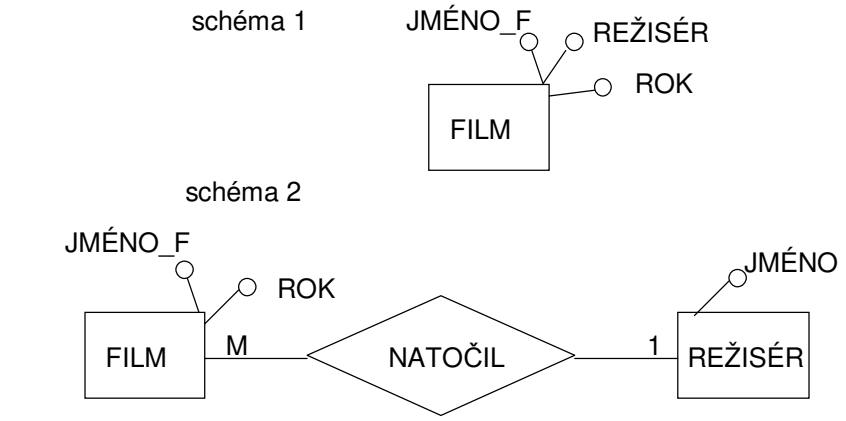
Integrace schémat - strukturální problémy

schéma 1

schéma 2

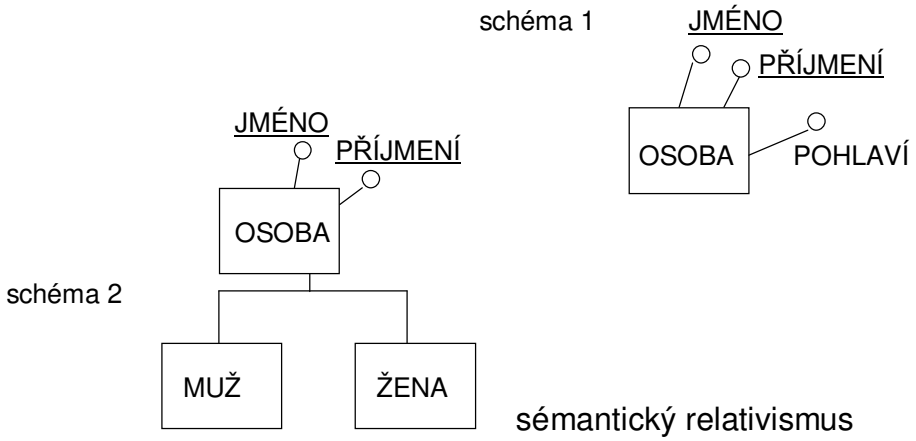
odlišná úroveň abstrakce

Integrace schémat - strukturální problémy



sémantický relativismus
v použití konstruktů

Integrace schémat - strukturální problémy



sémantický relativismus
v použití konstruktů

Analýza konfliktů ve jménech

Zdroje:

- *okolí pojmu* - pojmy dané přímými sousedy (*vlastnostmi pojmu*) odpovídajícího uzlu v E-R diagramu

  podobné pojmy mají větší množství společných vlastností a různá jména ($\Rightarrow$ synonyma)

  zmatené pojmy mají stejná jména a různé vlastnosti ($\Rightarrow$ homonyma)

Pz.: stejnost jmen a různost vlastností tvoří pro zmatené pojmy pouze postačující a nikoliv nutnou podmínku

- kardinality

Analýza konfliktů ve jménech

Modifikace schémat za účelem jejich integrace spočívají v přejmenovávání pojmů a případně k vytváření jejich nových vlastností.

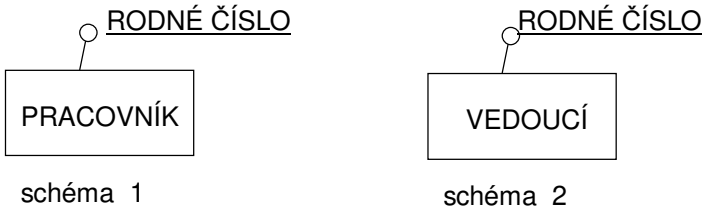
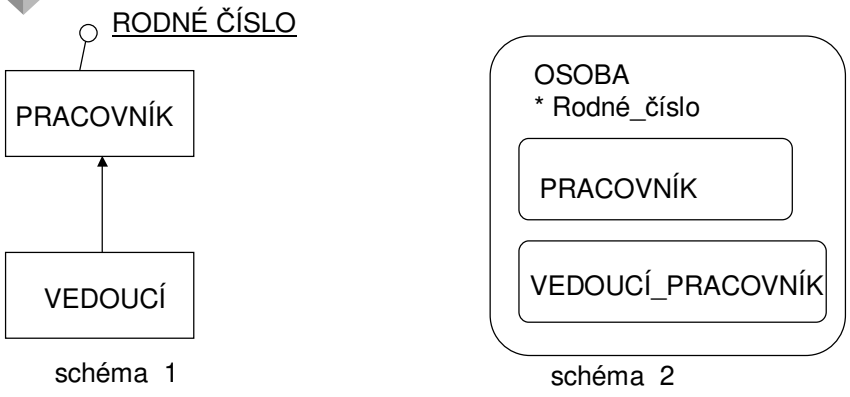


schéma 1 schéma 2

Analýza konfliktů ve jménech

řešení přejmenováním

Analýza konfliktů ve jménech



řešení hierarchií

Analýza konfliktů - kardinality

schéma 1

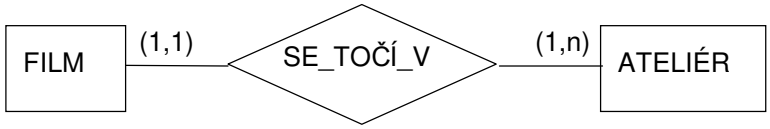
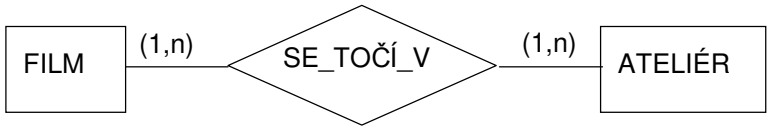


schéma 2



nekompatibilní specifikace

Analýza konfliktů – kardinality a typ členství

schéma 1

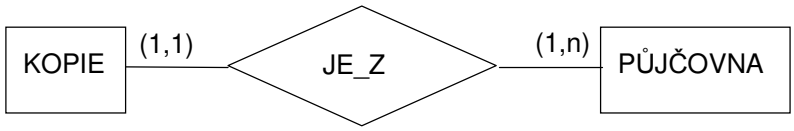
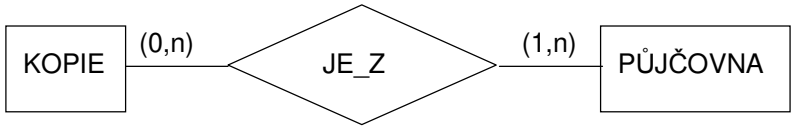
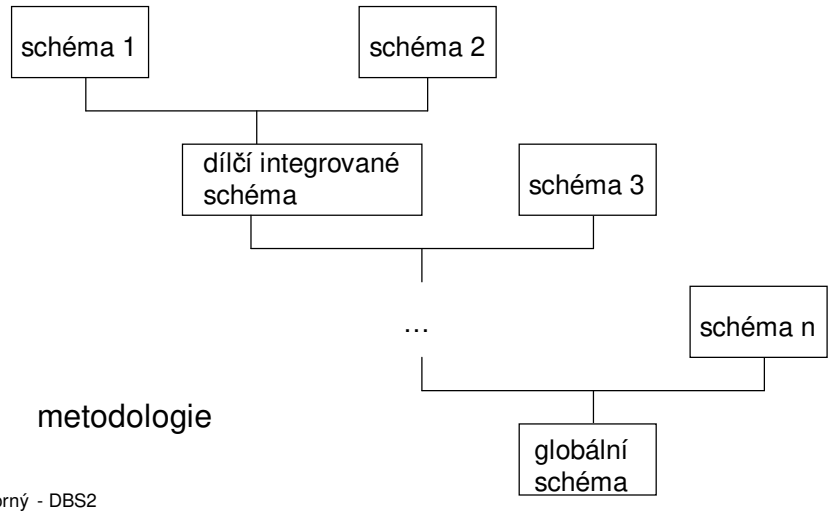


schéma 2



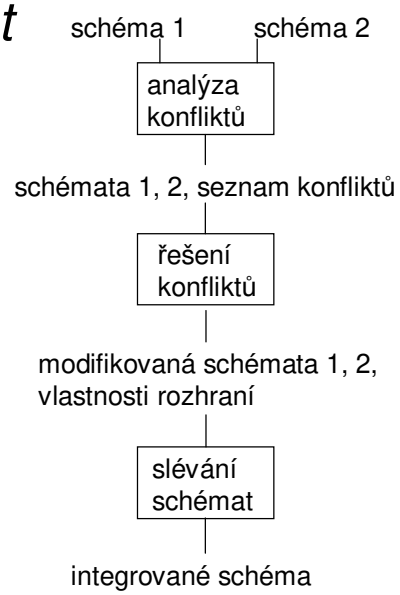
nekompatibilní specifikace

Integrace schémat



metodologie

Integrace schémat



metodologie

Příklad

Doména:

- publikační systém pracovníka katedry
- fakultní knihovna

Kroky:

- analýza obou dílčích domén
- analýza a řešení konfliktů
- modifikace schémat
- slévání schémat

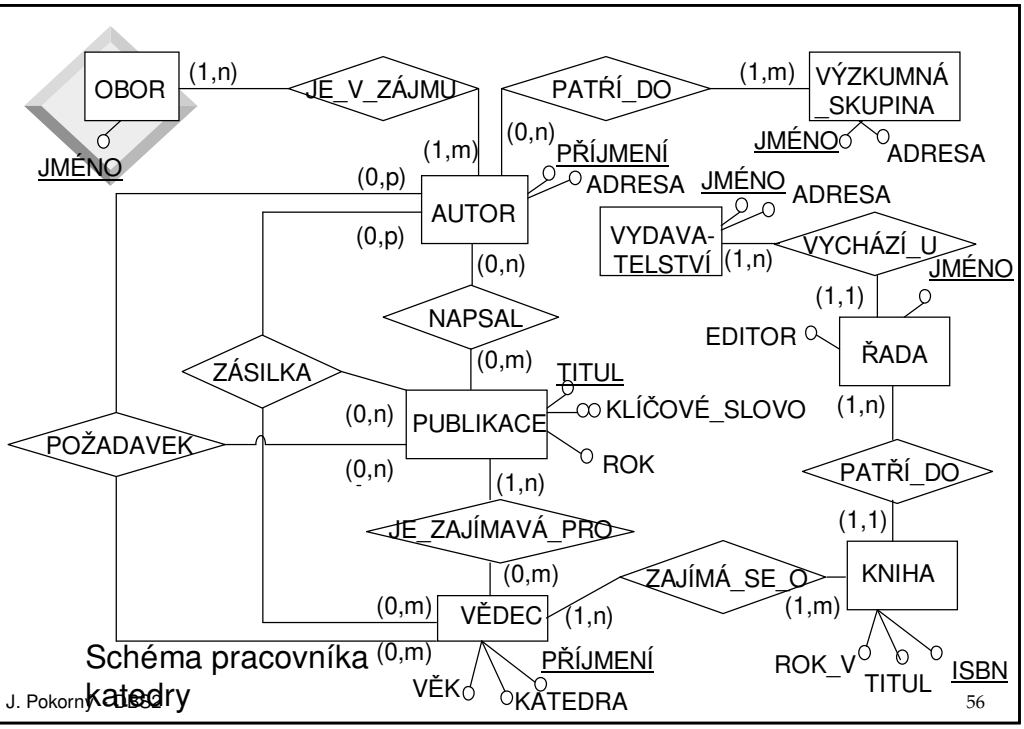
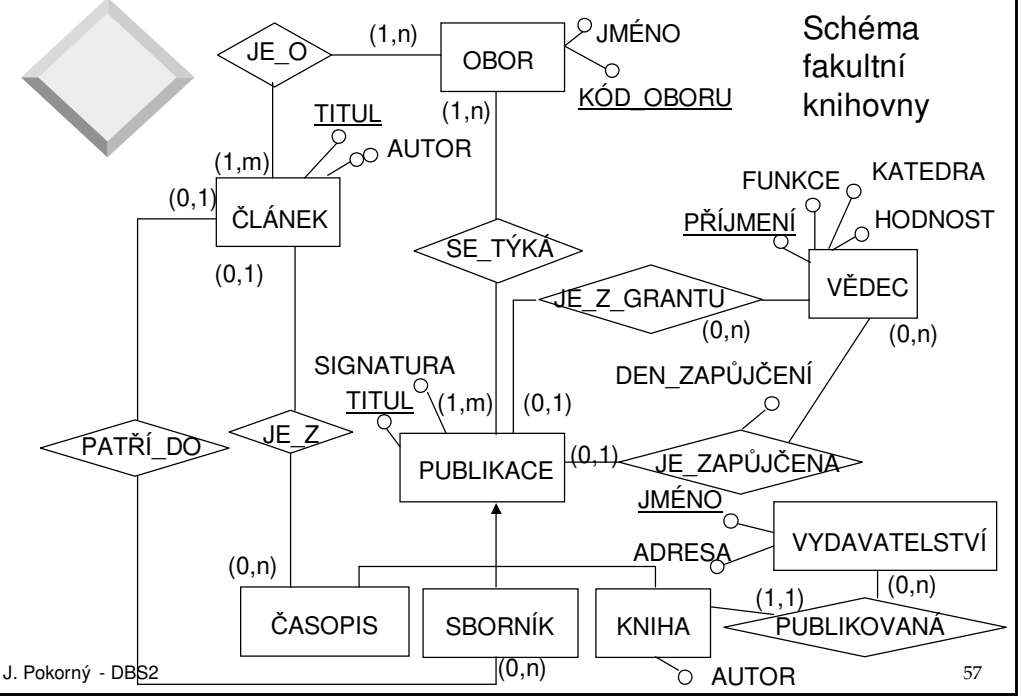


Schéma pracovníka katedry



Analýza konfliktů

- Typy entit OBOR a PUBLIKACE mají v obou schématech různé vlastnosti,
- KLÍČOVÉ_SLOVO je (vícehodnotovým) atributem typu entity PUBLIKACE.
- typ entity OBOR je vztažen k typu entity PUBLIKACE.
- VYDAVATELSTVÍ je v jednom schématu existenčně nezávislé, ve druhém ne.
- AUTOR je jednou existenčně nezávislý, jednou ne.

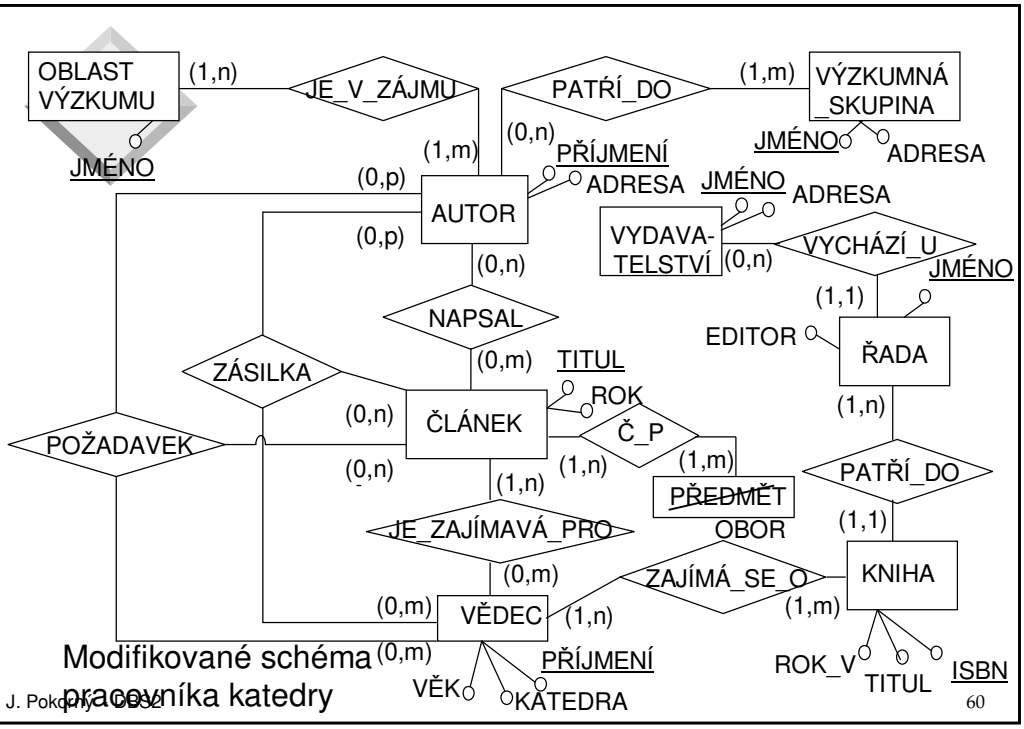
Řešení konfliktů

- Změny jmen typů entit, změny atributů na typy entit

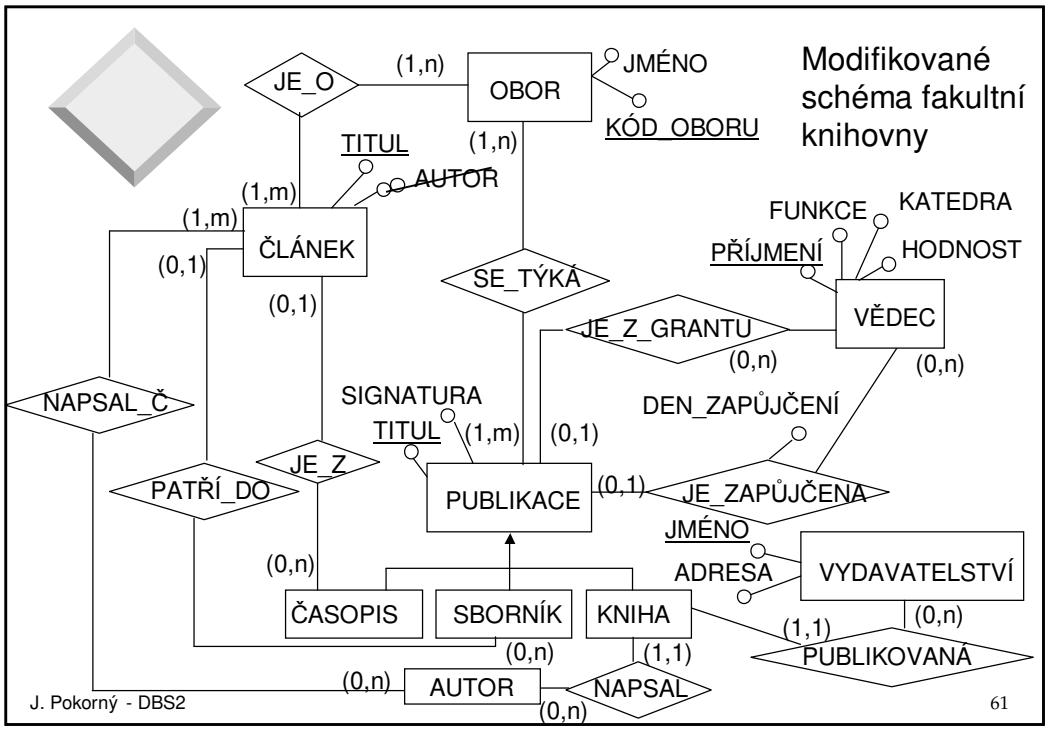
OBOR → OBLAST_VÝZKUMU

KLÍČOVÉ_SLOVO → PŘEDMĚT → OBOR

- komunikace s autory se týká ne obecných publikací, ale článků.
- PUBLIKACE → ČLÁNEK
- Rozdílné existenční závislosti ⇒ použije se “volnější” omezení, tj. min v min-max IO bude 0.



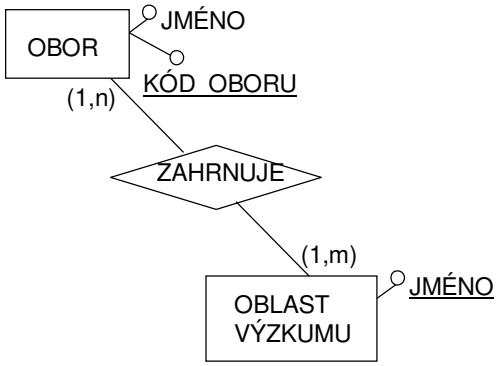
Modifikované schéma pracovníka katedry



Slévání modifikovaných schémat

- vznikne integrované schéma
- schéma obsahuje i něco „navíc“, např.

je nutný nový typ vztahu ZAHRNUJE



Funkční analýza

- modelování funkcí
- modelování událostí
- modelování toků dat

Hierarchie funkcí

- (pracovní) funkce
  - používané identifikátory

    Př.: Založ nový požadavek na rezervaci

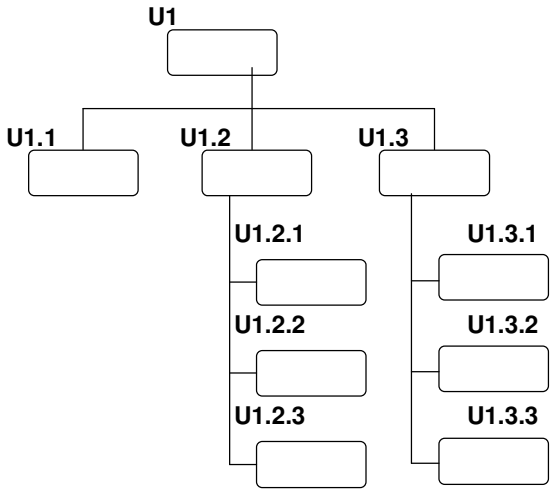
identifikuj	sestav	změň
kontroluj	spoj	vytvoř
proveď	testuj	zaúčtuj
odešli	urči	zruš
odstraň	zkontroluj	vyber

  - z popisu funkcí se vylučují informace typu *KDO*, *KDE*, *KDY* a *JAK*.

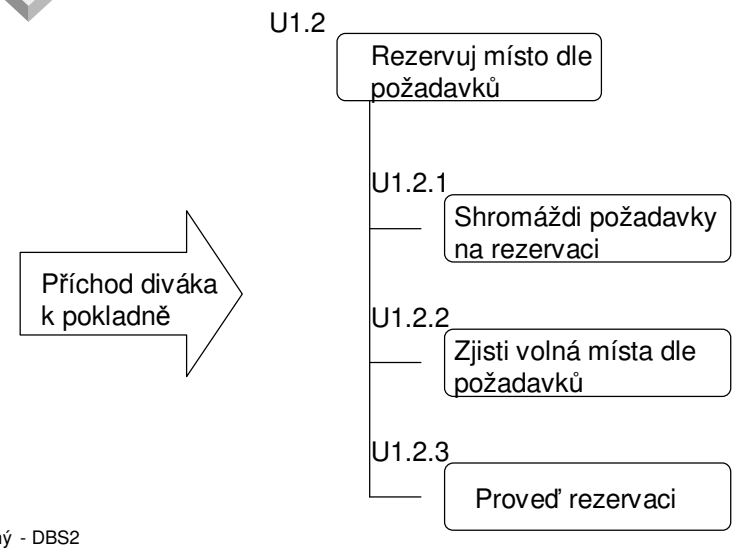
Hierarchie funkcí - notace

U13

Testuj požadavek na výdej lístku



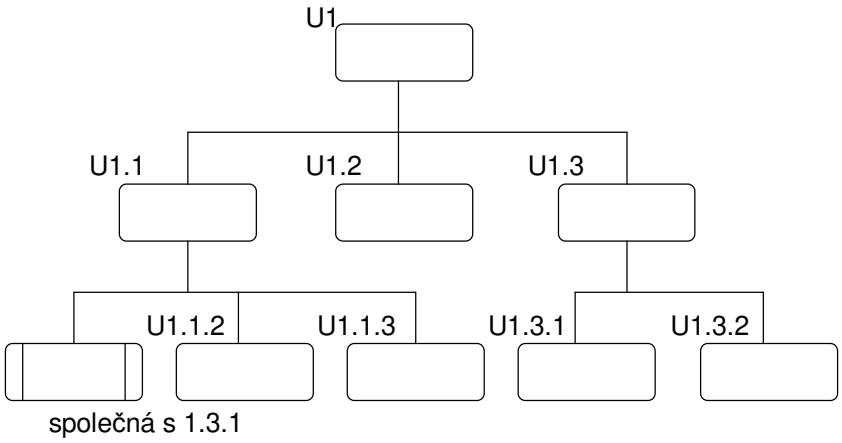
Hierarchie funkcí - triggery



Typy funkcí

- *atomické* (listové v hierarchii)
- *elementární* (nemusí být atomické, podobné transakcím)
- *společné* (jedna z nich je *hlavní*)

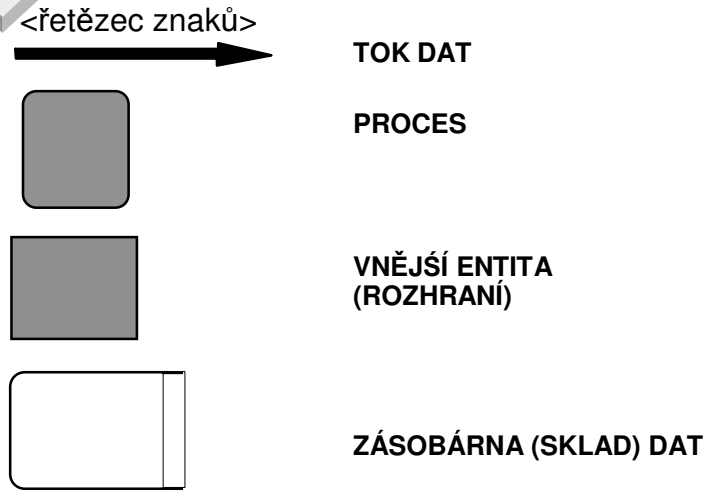
Typy funkcí



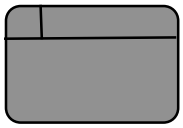
Použití funkcí k funkční analýze

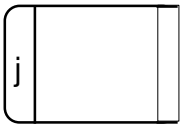
- Metoda: interview
- Kvalita schématu funkcí:
  - přesné a na každé úrovni srozumitelným způsobem čitelné,
  - úplné,
  - stručné ve výrazivu,
  - vyvážené tak, že je snadné ho číst.
- Vztah k E-R
  - existují funkce, které zmiňují typy entit použité v E-R schématu,
  - existují funkce, které vytvářejí, mění, odstraňují entity, vztahy, hodnoty atributů typů daných v E-R schématu.

Diagramy toků dat



Diagramy toků dat



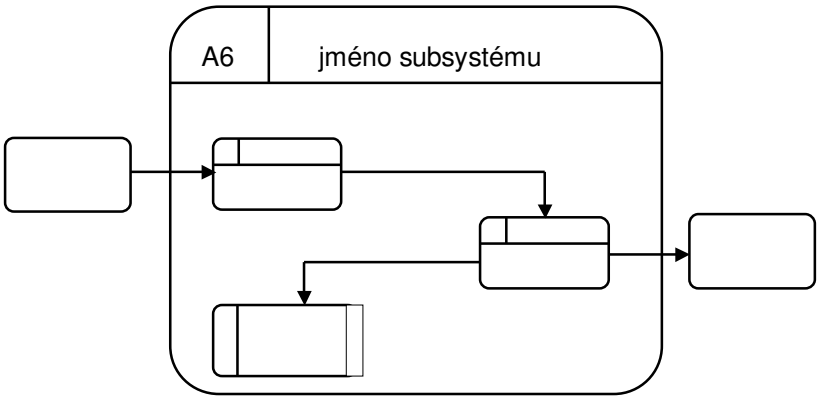


→	rozhraní	proces	zásobárna
rozhraní	**NE**	**ANO**	**ANO***
proces	**ANO**	**ANO**	**ANO**
zásobárna	**ANO***	**ANO**	**NE**

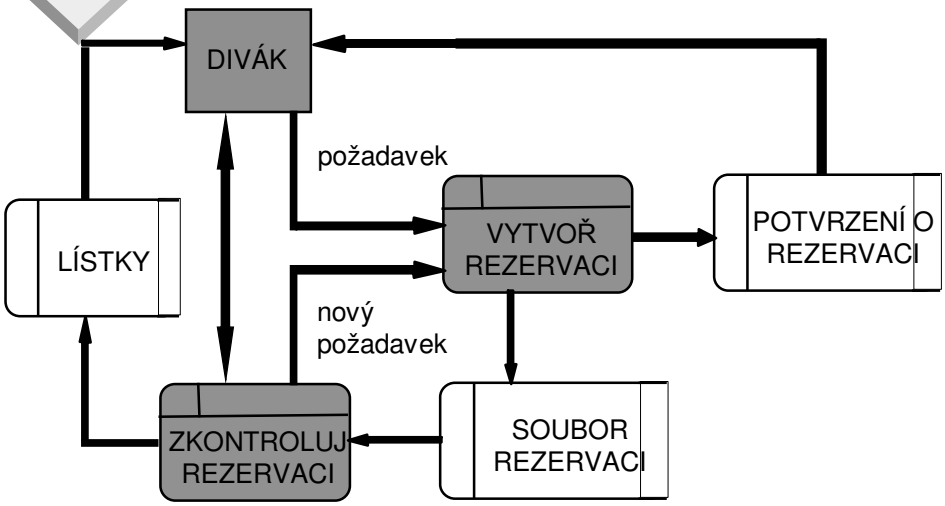
Možnosti toků dat

proces pokrývá několik pracovních funkcí

Diagramy toků dat – hnízdění procesů

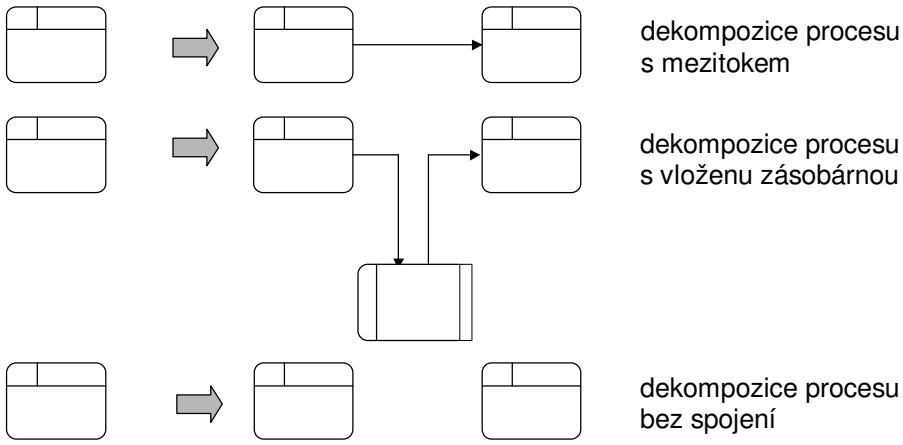


Diagramy toků dat



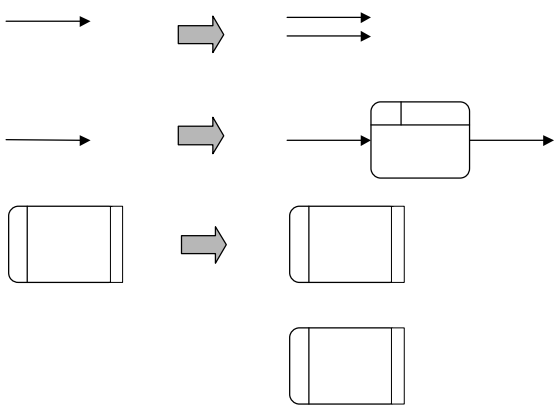
Rezervace lístků – diagram toků dat

Diagramy toků dat



Elementární transformace

Diagramy toků dat

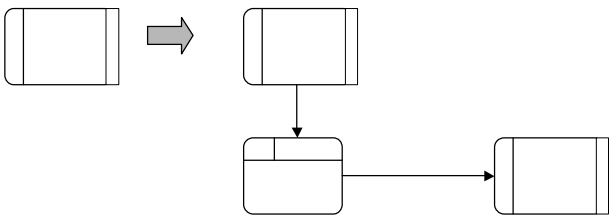


dekompozice toku dat

zjemnění toku dat

dekompozice zásobárny

Diagramy toků dat

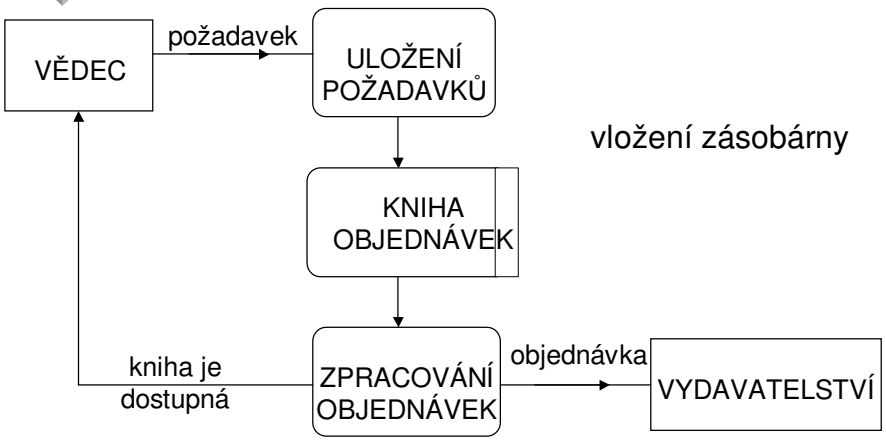


dekompozice zásobárny s vloženým procesem

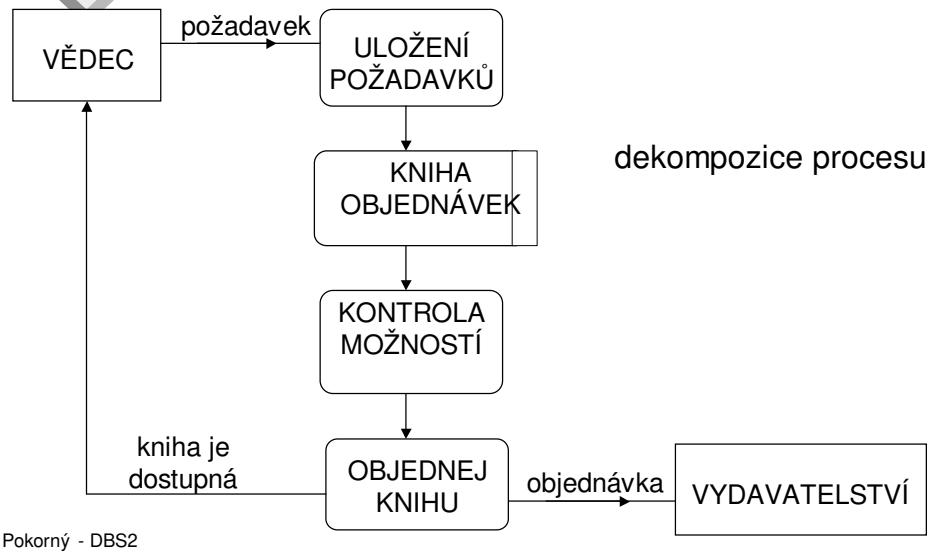
Strategie návrhu shora dolů

Skeletální schéma

Strategie návrhu shora dolů - 2. stadium

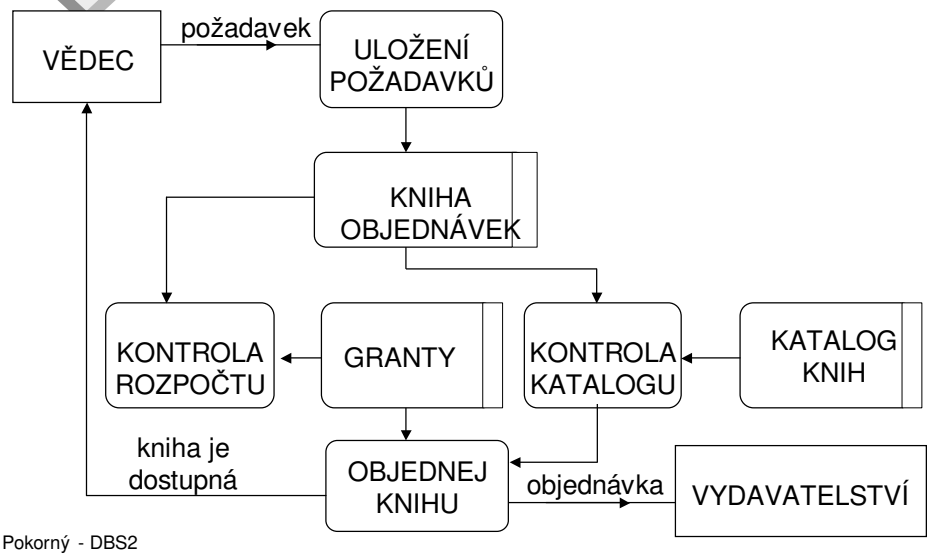


Strategie návrhu shora dolů - 3. stadium



dekompozice procesu

Strategie návrhu shora dolů - 4. stadium

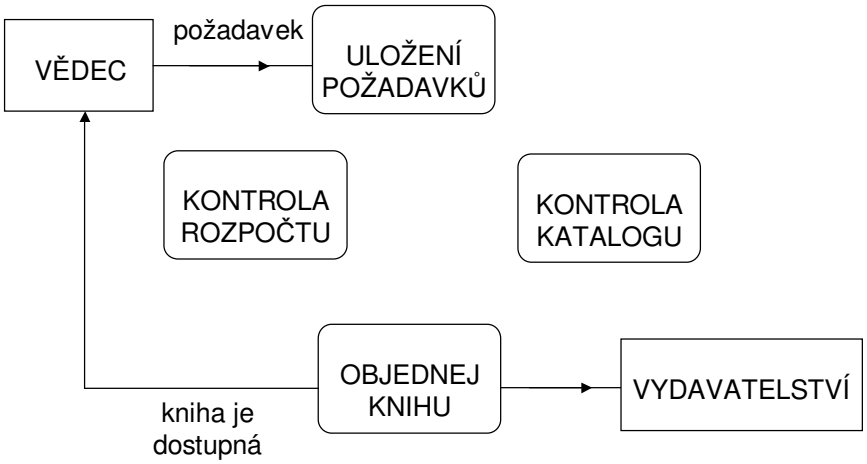


Strategie návrhu zdola nahoru

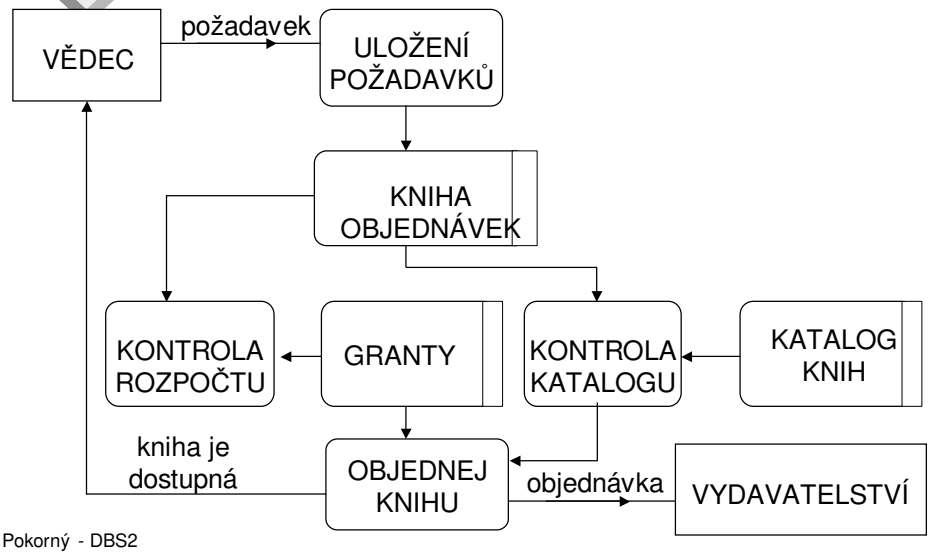
Kroky:

1. zavedou se všechny elementární procesy,
2. zavedou se rozhraní,
3. procesy se spojí toky, zavedou se zásobárny.

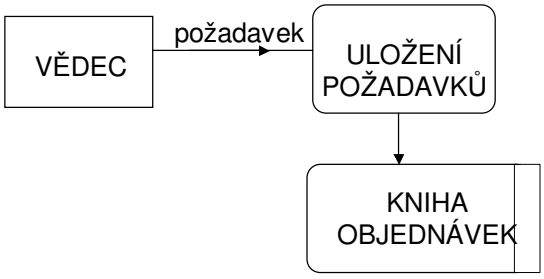
Strategie návrhu zdola nahoru - 2. stadium



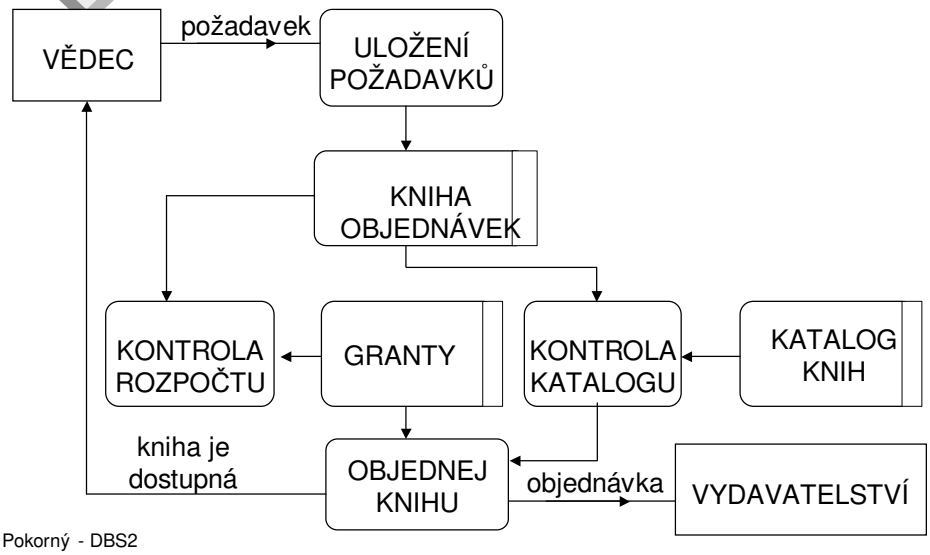
Strategie návrhu zdola nahoru - 3. stadium



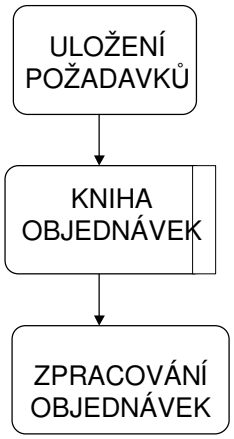
Strategie návrhu zevnitř ven



Strategie návrhu zevnitř ven - výsledek



Smíšená strategie

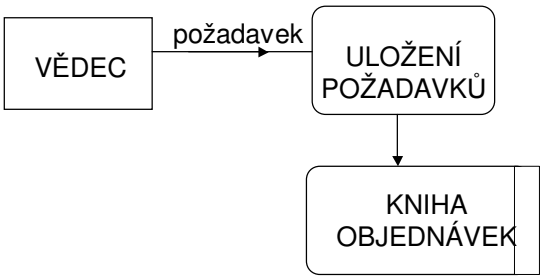


Kroky:

1. zvolí se skeletální schéma
2. rozpracují se jeho části (event. různými strategiemi)

Smíšená strategie - 2a. stadium

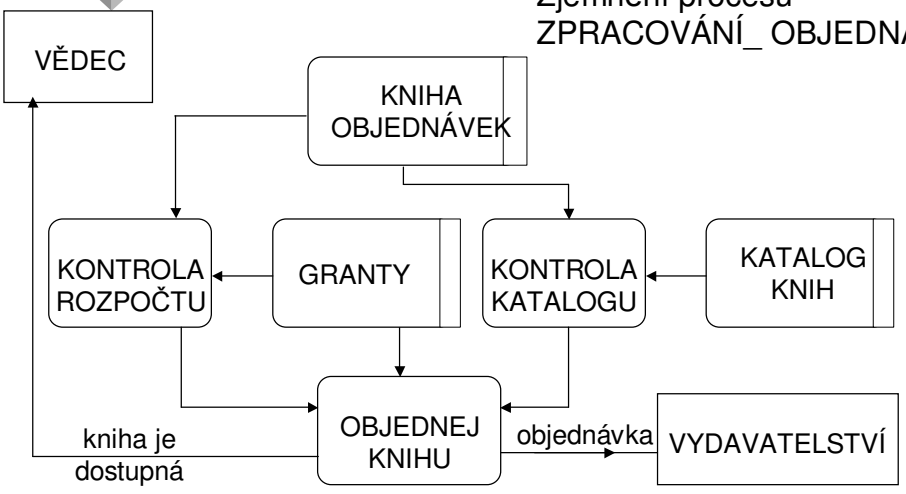
Pz.: vše shora dolů



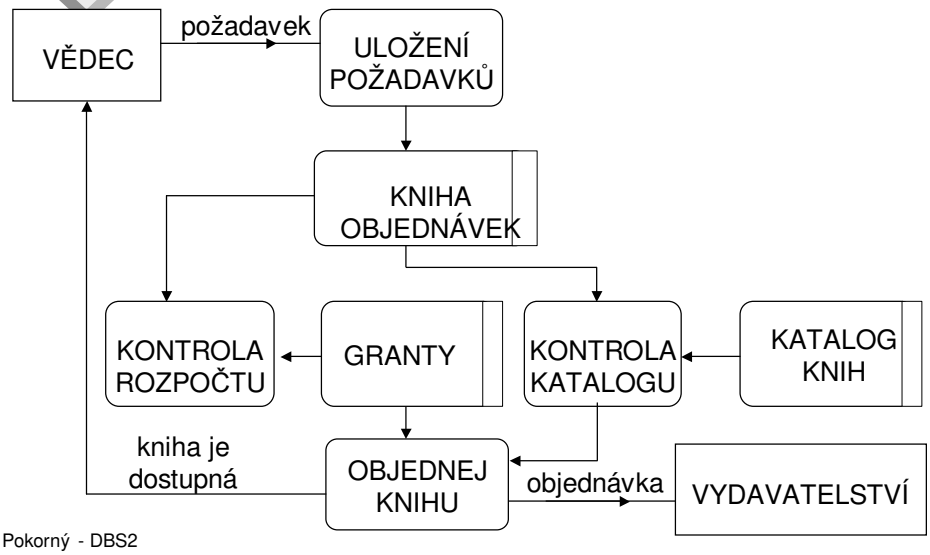
Zjemnění procesu
ULOŽENÍ_POŽADAVKŮ

Smíšená strategie - 2b. stadium

Zjemnění procesu
ZPRACOVÁNÍ_ OBJEDNÁVEK



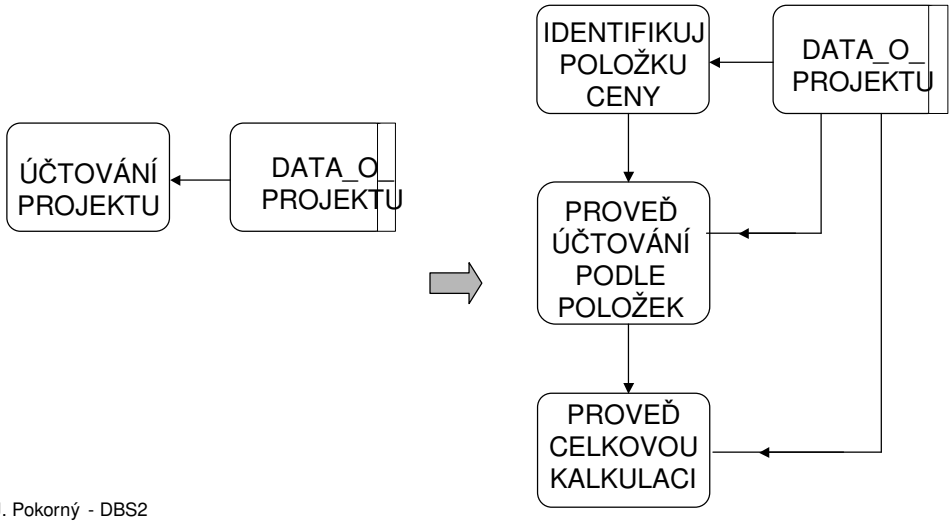
Smíšená strategie - výsledek



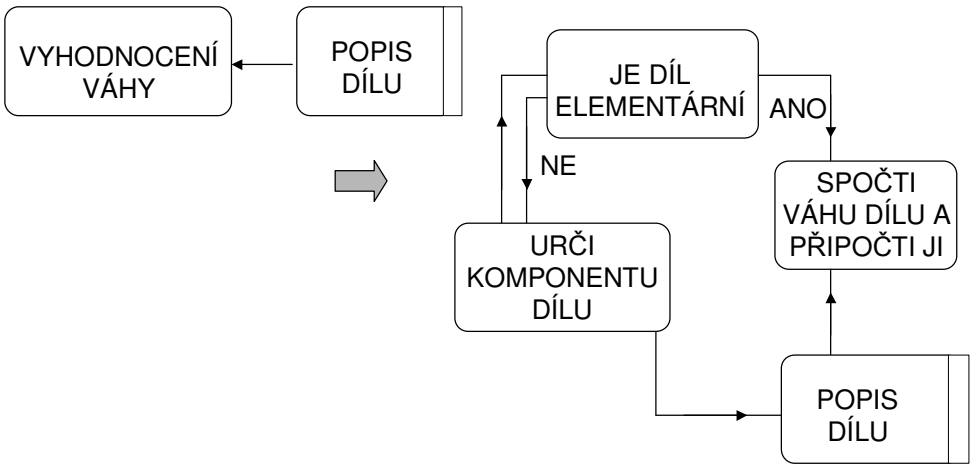
Obecné zásady pro vytváření diagramu toků dat

- zaměřit se na CO specifikujeme, a nepopisovat JAK mají procesy probíhat
- nezjemňovat procesy, ztrácí-li se již jejich globální struktura

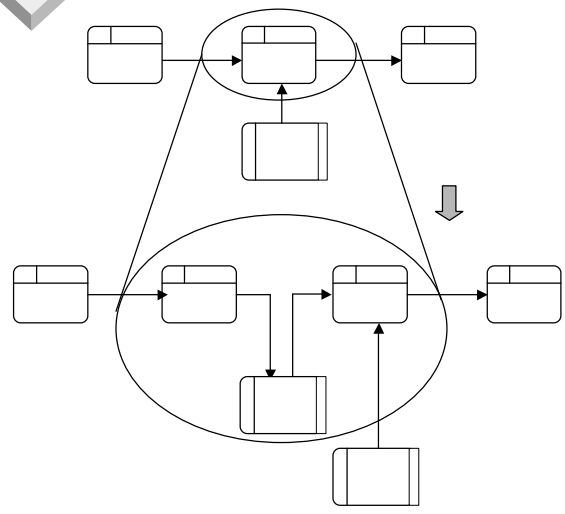
Příklad procedurality (1)



Příklad procedurality (2)



Vlastnosti diagramů toků dat

- funkční nezávislost
  - separabilita
  - jednoduchost integrace
  - flexibilita
- úplnost
- korektnost
- čitelnost
- minimalita

Korektnost - ekvivalence hranic