

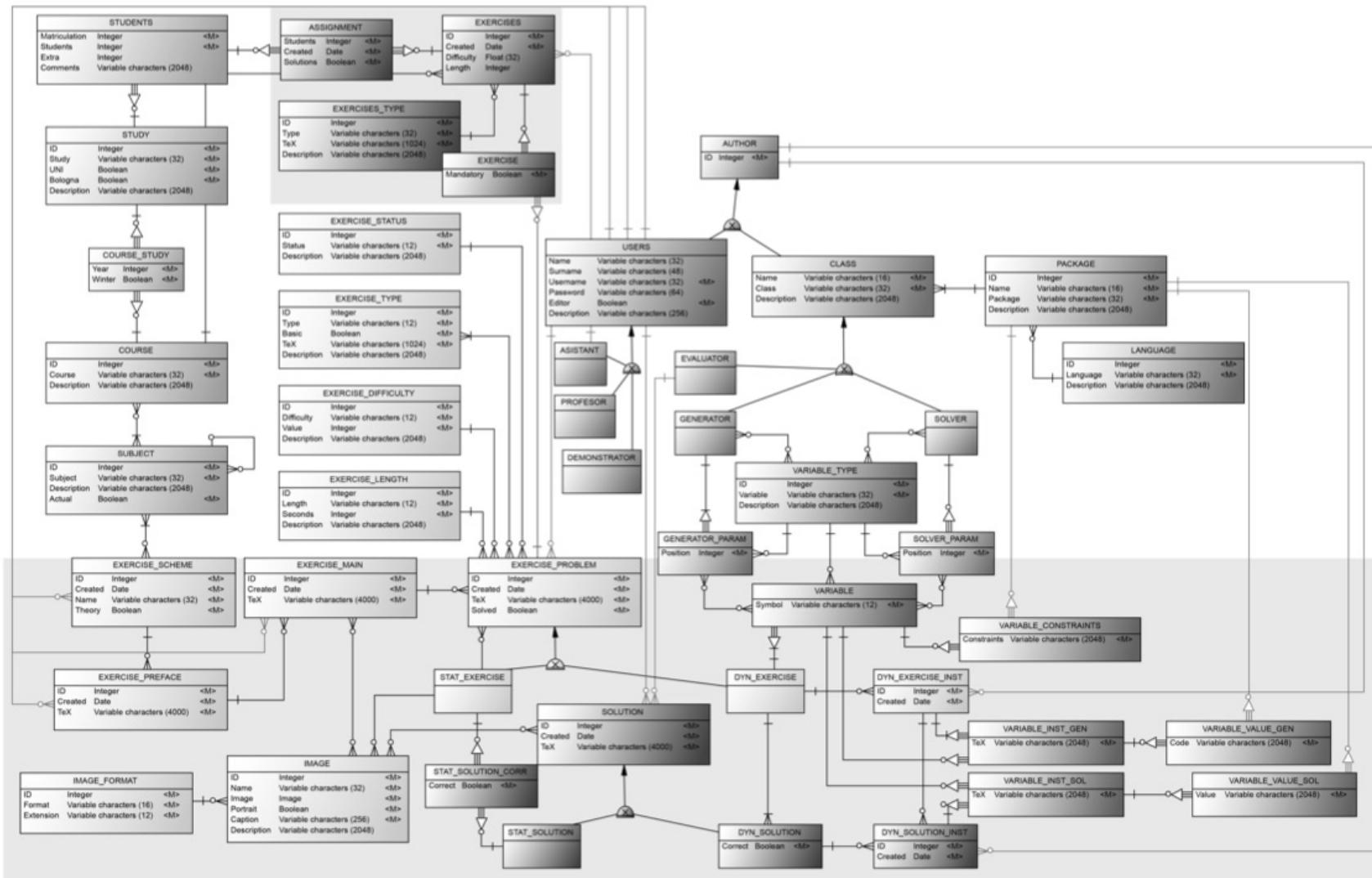
Indeksiranje podatkovnih baz

Osnove podatkovnih baz

2. letnik univerzitetnega študija na FRI

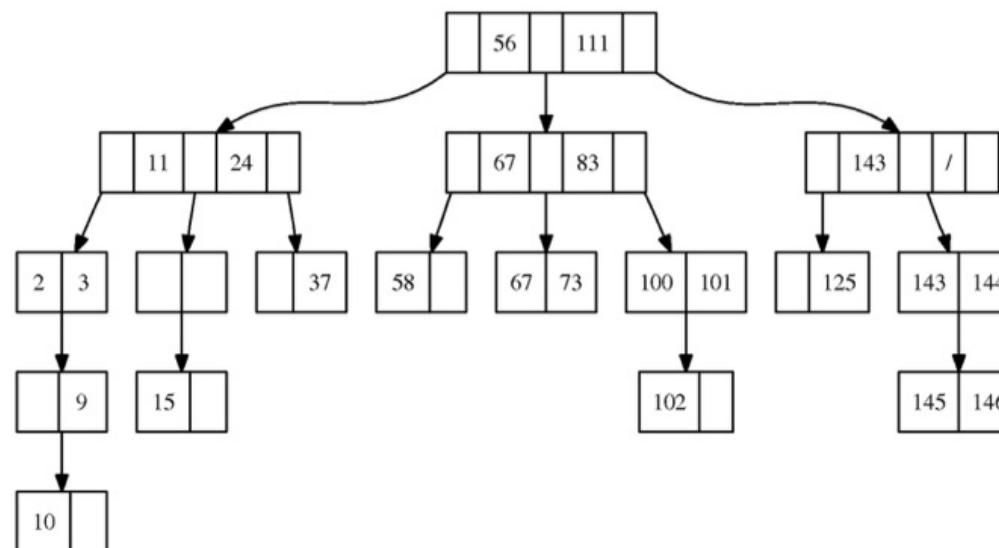
2021/22

Konceptualni model bLUE



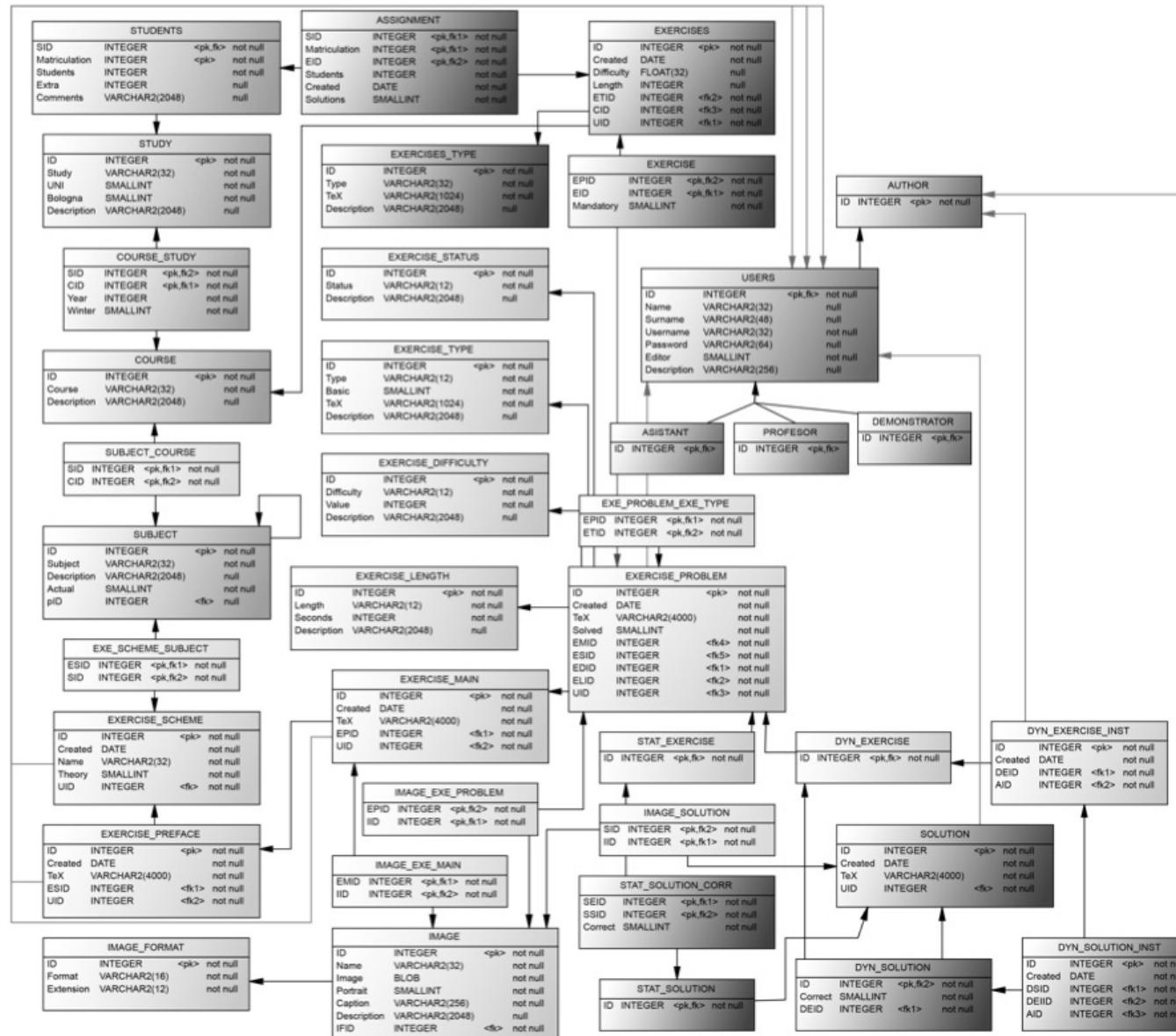
ISAM indeks

- ▶ ISAM indeks je statično (ne)uravnoteženo drevo
 - ▶ prelivne strani se prilagajajo pri posodabljanu zapisov
 - ▶ časovna zahtevnost iskanja zapisov $\approx O(\log n) \ll O(n)$



ISAM indeks za namišljeni atribut

Logični model bLUE



1. naloga Prodajalna avtomobilov

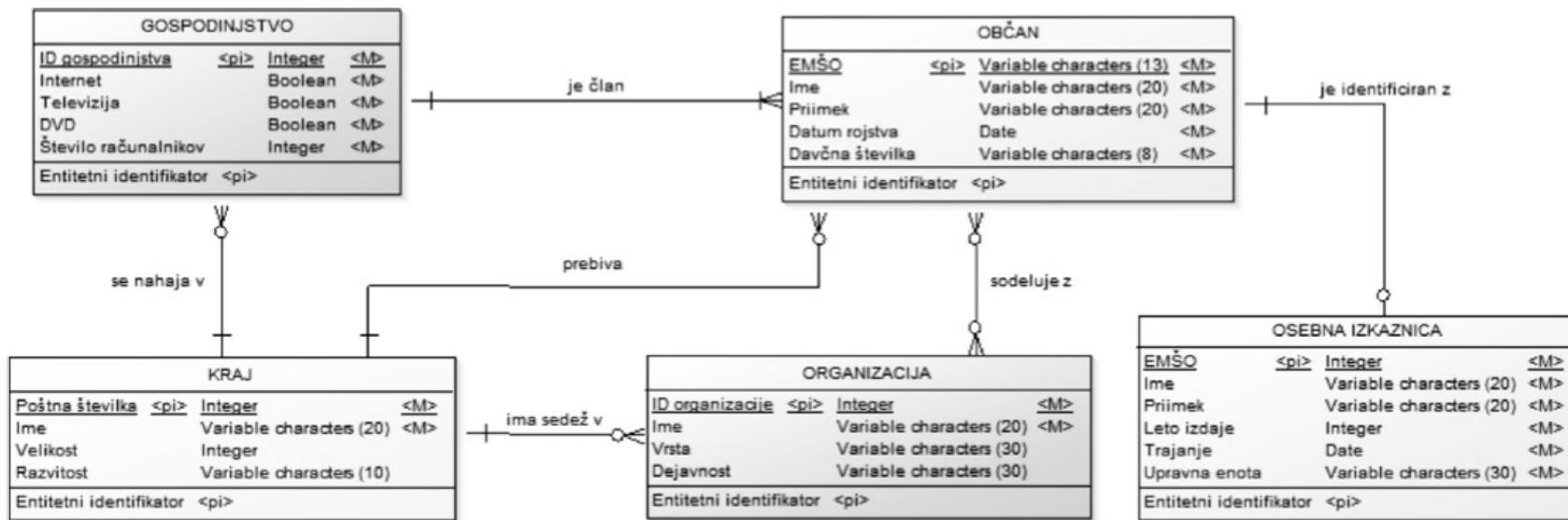
Študenti pri predmetu OPB v svojem prostem času vodijo prodajalno luksuznih avtomobilov, da si lahko v poletnih mesecih privoščijo zaslužen dopust.

Pri ceniku avtomobilov v prodajalni avtomobilov upoštevajo, da se cene avtomobilov lahko spreminja v odvisnosti od datuma in raznih akcij.

Cenik vsebuje naslednje podatke: termin, za katerega cenik velja (lahko začetni in končni datum ali samo datum začetka veljave), morebitno ime akcije in postavke cenika.

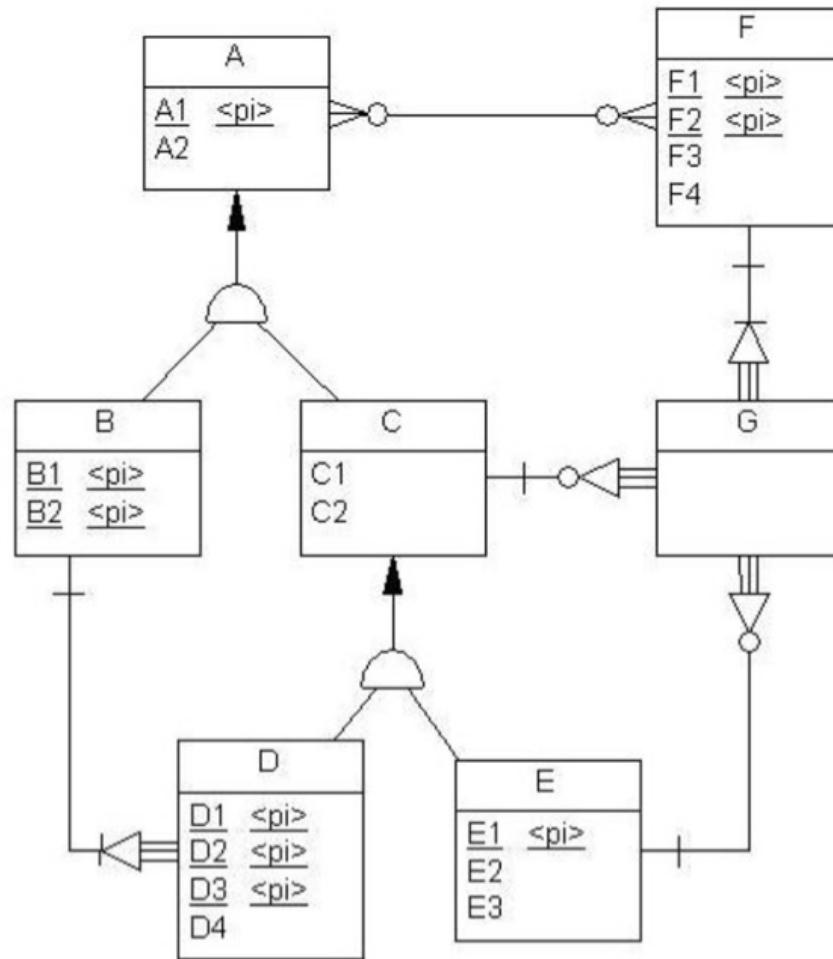
V vsaki postavki cenika je naveden artikel (avtomobil), cena brez DDV ali cena z DDV ter morebitni popust ob akciji.

1. naloga Gospodinjstvo

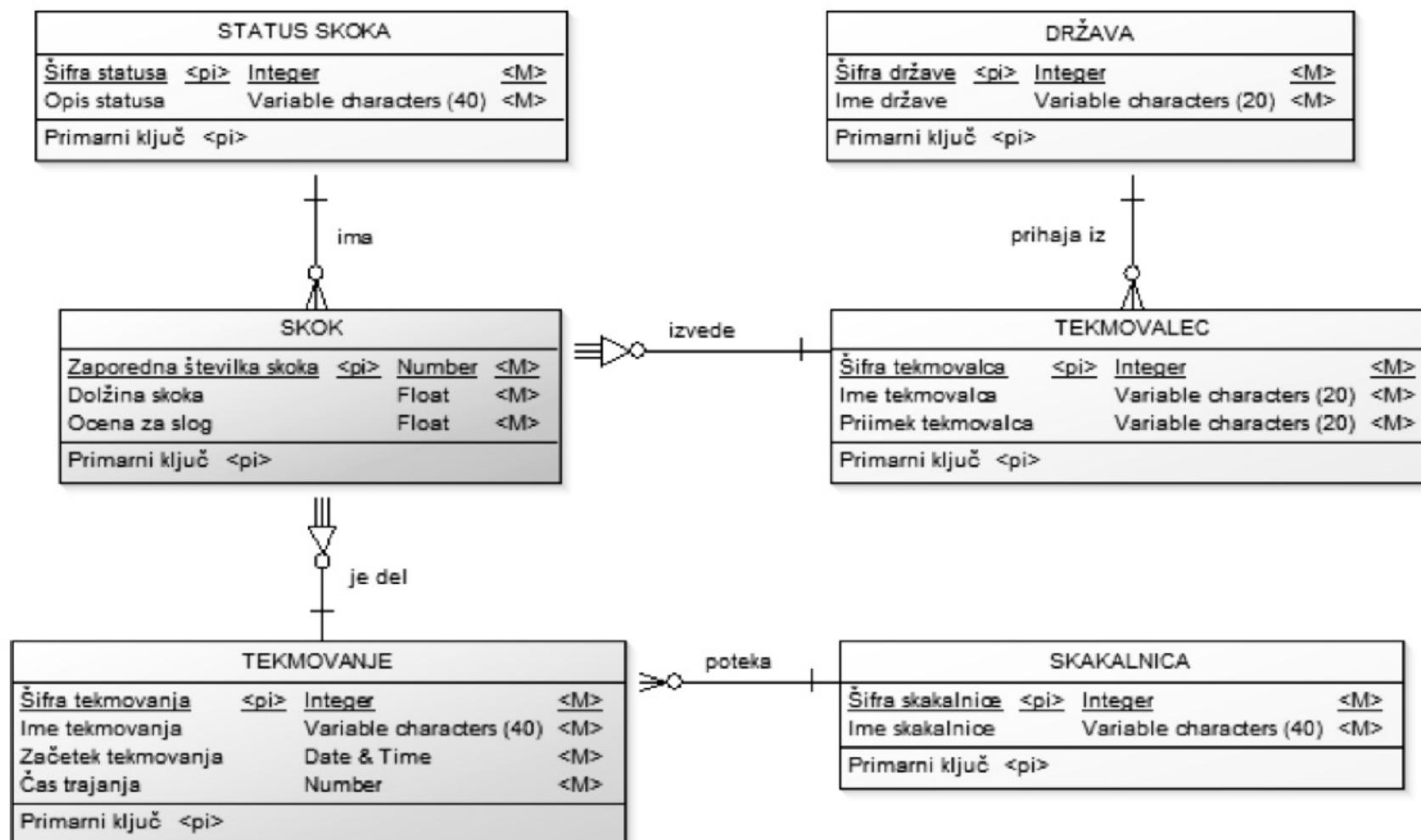


4. naloga ABC

Pri preslikavi specializacije tvorite tri tabele tako, da se iz nadrejene v podrejene tabele prenesejo le atributi, ki so del primarnega ključa.

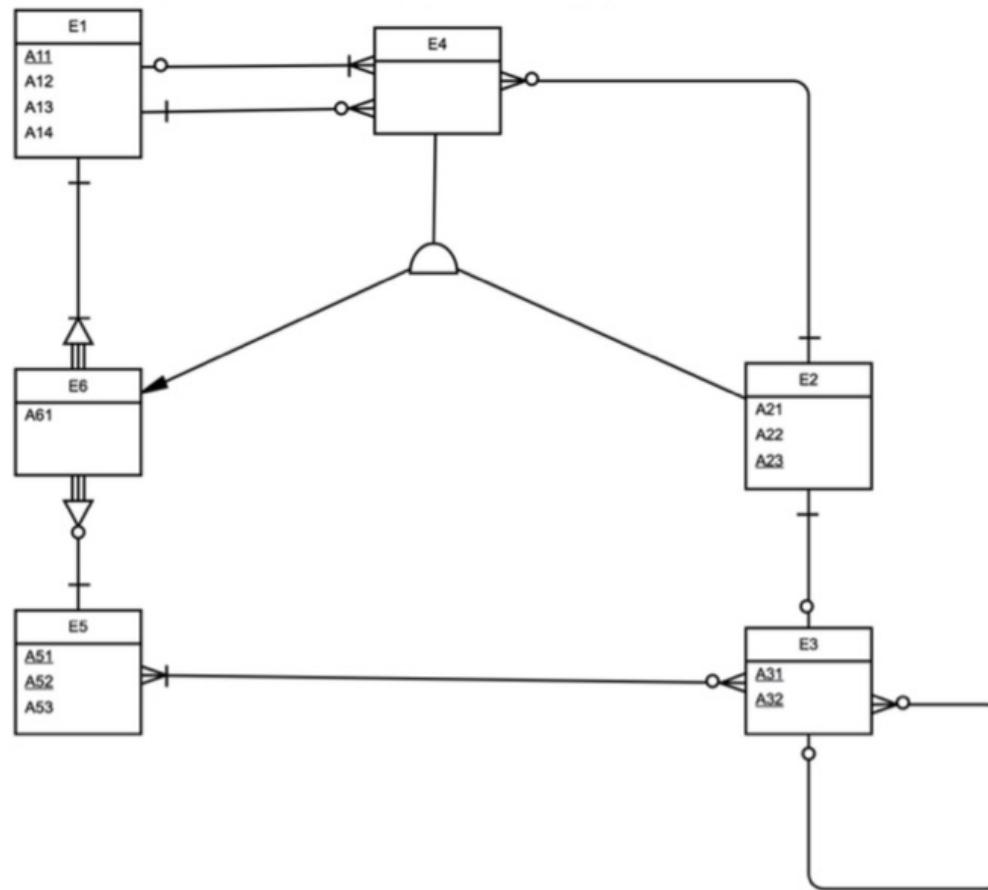


2. naloga Smučarski skoki



5. naloga 123

Pri preslikavi specializacije tvorite tri tabele tako, da se iz nadrejene v podrejene tabele prenesejo le atributi, ki so del primarnega ključa.



$E5(\underline{A51}, \underline{A52}, A53)$

$E3(\underline{A31}, \underline{A32}, \#E2-A23, \#E2-A11, \#E2-A51, \#E2-A52, \#E3-A31, \#E3-A52)$

$E5-E3(\#A51, \#A52, \#A51, \#A52)$

$E1(A11, A12, A13, A14)$

$E6(A61, \#A11), \#A51, \#A52)$

$E2(A21, A22, A23), \#A11, \#A51, \#A52)$

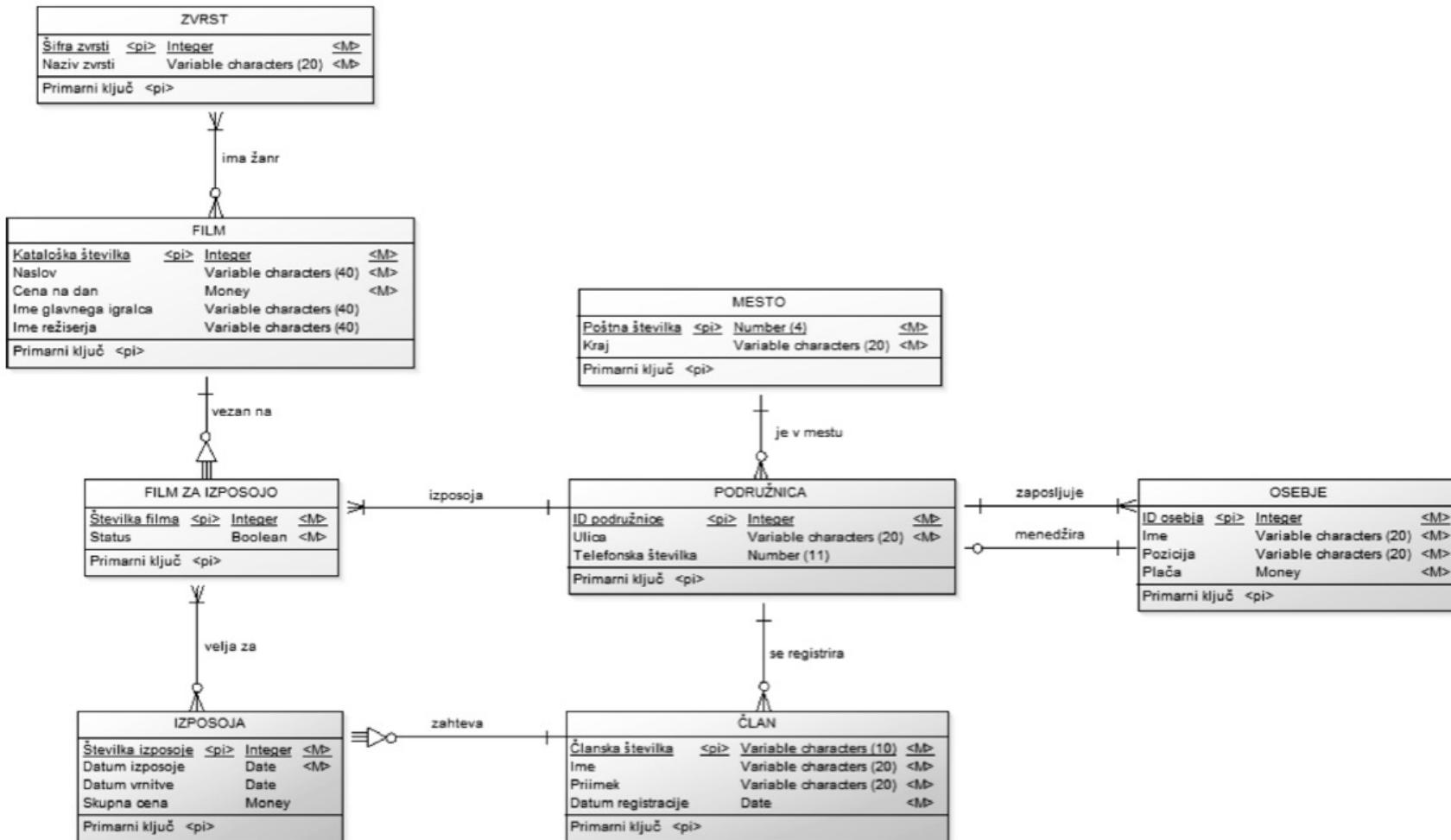
$E4(\#A11, \#A51, \#A52, \#E1-A11-1, \#E1-A11-2, \#E2-A11, \#E2-A51, \#E2-A52, \#A23)$

3. naloga Videoteka Pirat

... Vsak film je identificiran s kataloško številko, dočim pa lahko obstaja več kopij istega filma v posamezni podružnici, od katerih je vsaka identificirana s svojo številko.

Preden si stranka lahko izposodi film, se mora registrirati kot član v svoji lokalni podružnici. Vsakemu članu se dodeli članska številka, ki je unikatna v okviru vseh podružnic. Ko se stranka registrira, lahko začne z izposojo filmov. V okviru ene izposoje si lahko izposodi največ 10 filmov. Podatki, ki se hranijo o izposoji, so naslednji: številka izposoje, polno ime in številka člana, številke in naslovi filmov, skupna cena in datum izposoje ter vrnitve.

3. naloga Videoteka Pirat



4. naloga Fakulteta

V fakultetni podatkovni bazi so shranjeni podatki o študentih in učiteljih. Študentje so predstavljeni z imenom in priimkom, starostjo, spolom, rojstnim krajem ter s podatki o mestih stalnega prebivališča in obdobjih, ki so jih v njih preživeli. Poleg tega za vsakega študenta hrаниmo tudi podatke o opravljenih izpitih (predmet, ocena, nosilec, datum) ter katere predmete obiskuje v tekočem študijskem letu (vključno z imenom predavatelja), za absolvente pa tudi naslov diplomskega dela in ime mentorja.

V podatkovni bazi je shranjen tudi urnik, ki pove katerega dne, kje in kdaj se začnejo predavanja posameznega predmeta. Vsak predmet se predava le enkrat tedensko in v tekočem letu ga predava le en učitelj. Učitelji so v podatkovni bazi predstavljeni z imenom in priimkom, starostjo, nazivom, raziskovalnimi področji, s katerimi se ukvarjajo, ter katedro, katere člani so. Za učitelje in katedre so zapisane tudi telefonske številke, kjer so dosegljivi.

4. naloga WikiLeaks

embassy (id, name)

privacy (id, classification)

cable (id, date, header, content, #privacy_id, #embassy_id)

reference (#cable_id, #ref_cable_id)

- (a) Pojasnite čas izvajanja naslednje poizvedbe

*SELECT SQL_NO_CACHE * FROM wiki WHERE DATE(date) = '2009-12-10';*

\ če posliši shoci funkcijo base ne bo uporabljala metsov

- (b) Ustvarite skupen indeks za atributa *date* in *name*

INDEX (date, name)

- (c) Poskusite ustvariti enoličen indeks za atribut *name*

ALTER TABLE wiki ADD UNIQUE INDEX (name)

\ preprečičev duplikatov

5. naloga WikiLeaks

embassy (id, name)

privacy (id, classification)

cable (id, date, header, content, #privacy_id, #embassy_id)

reference (#cable_id, #ref_cable_id)

- (a) Ustvarite tabelo *subwiki* ter ji določite primarni ključ

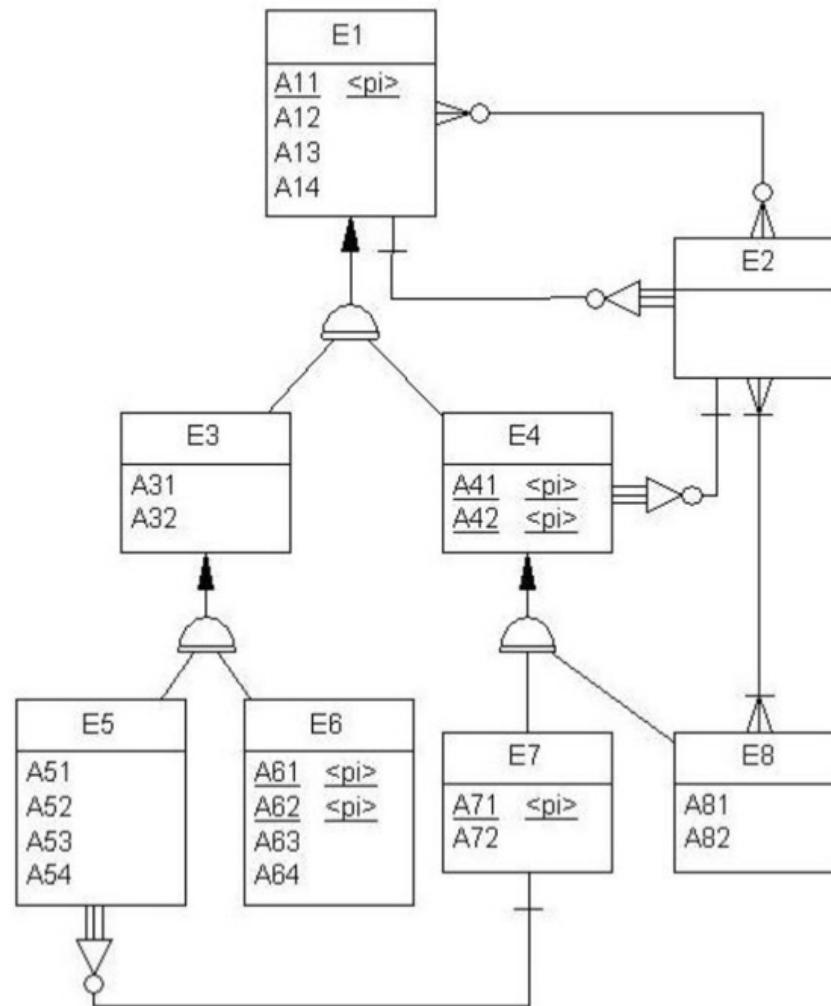
*CREATE TABLE subwiki AS SELECT * FROM wiki WHERE date < '1975-12-31' ORDER BY date;*

- (b) Narišite bitna indeksa za atributa *name* in *classification*
-

- (c) Kometirajte smiselnost bitnega indeksa za atribut *id*
-

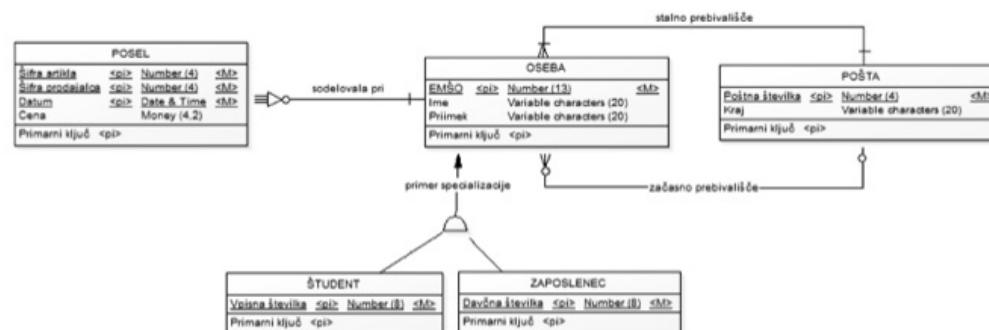
6. naloga 123

Pri preslikavi specializacije tvorite tri tabele tako, da se iz nadrejene v podrejene tabele prenesejo le atributi, ki so del primarnega ključa.



Koncepti logičnega načrtovanja

Konceptualni model	Logični model
entitetni tip	→ tabela
entiteta	→ zapis v tabeli
atribut	→ stolpec v tabeli
entitetni identifikator	→ primarni ključ
povezava 1:1	→ združimo v eno tabelo
povezava 1:0	→ tuji ključ v podrejeni tabeli
povezava 1:n	→ tuji ključ v podrejeni tabeli
povezava m:n	→ nova tabela s tujima ključema
šibka entiteta	→ tuji ključ del primarnega ključa
rekurzivna povezava	→ tuji ključ ali nova tabela
specializacija	→ nadrejena in/ali podrejene tabele → vsi atributi ali le primarni ključ

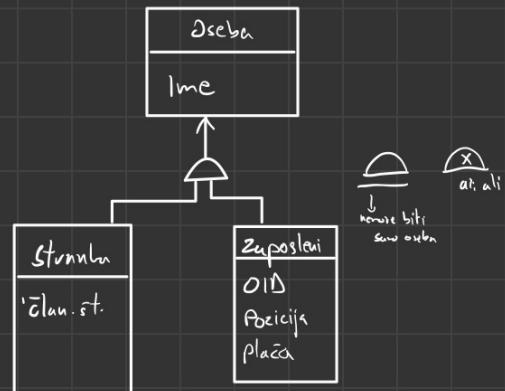
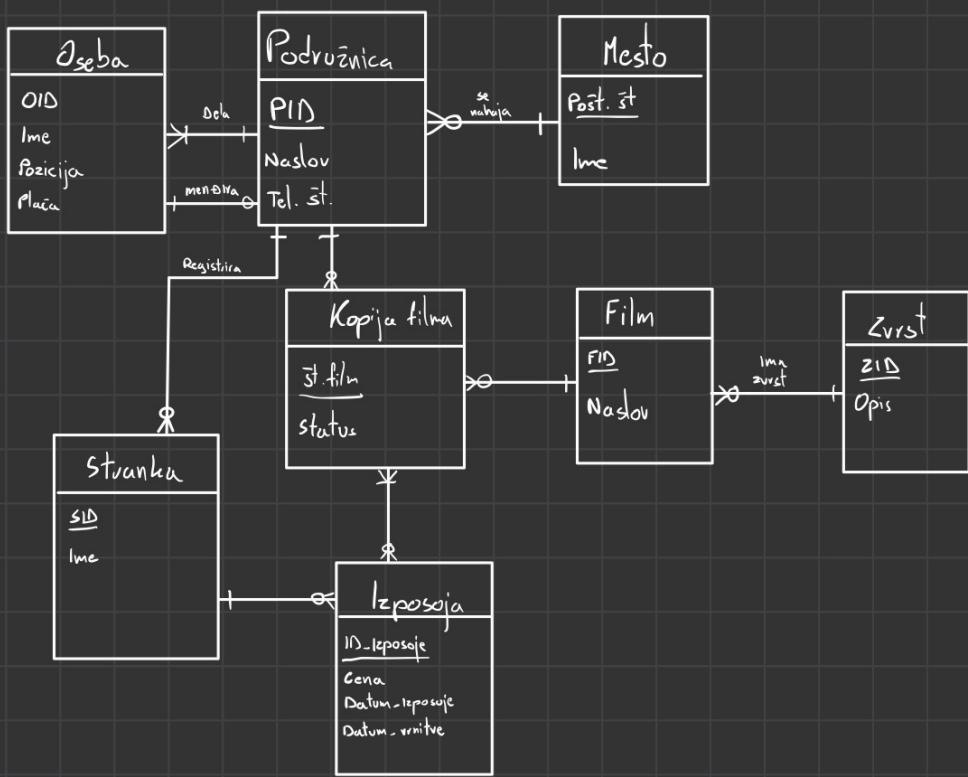


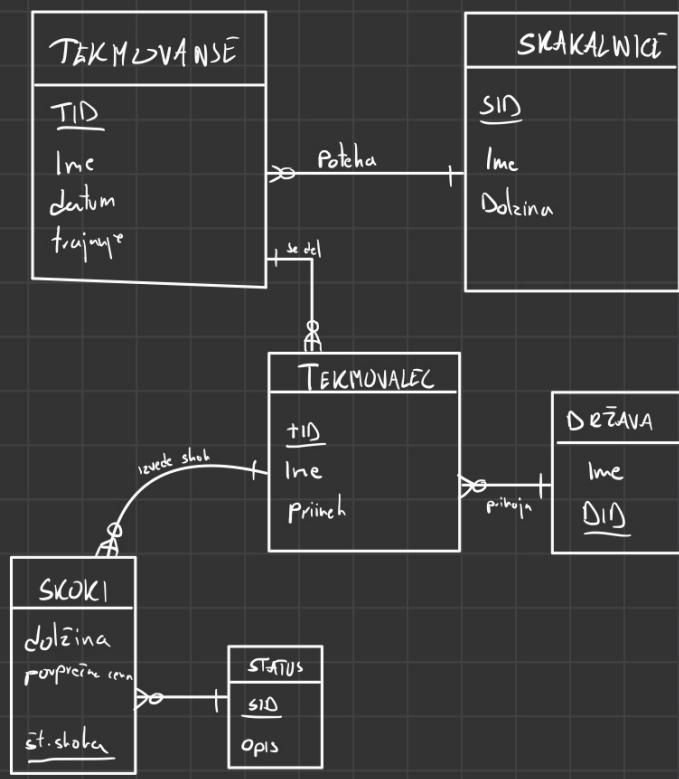
Logični podatkovni model

Osnove podatkovnih baz

2. letnik univerzitetnega študija na FRI

2021/22





2. naloga Smučarski skoki

V Planici in drugod po svetu vsako leto prirejajo tekmovanja v smučarskih skokih in poletih.

O tekmovanjih zbiramo naslednje podatke: datum in čas začetka tekmovanja, ime tekmovanja (npr. 20. tekmovanje za svetovni pokal), predviden čas trajanja in na kateri skakalnici poteka (npr. 90 metrska skakalnica).

Posamezno tekmovanje lahko poteka le na eni skakalnici. Na tekmovanje se lahko prijavijo tekmovalci, za katere moramo poznati ime, priimek in ime države iz katere prihajajo (npr. Norveška).

Za vsakega tekmovalca vodimo tudi podatke o skokih, ki jih je izvedel na posameznem tekmovanju. Za vsak skok poznamo dolžino skoka v metrih (npr. 158,6 m), povprečno oceno za slog v točkah (npr. 18,5), status skoka (npr. uspešen, razveljavljen, padec ipd.) in zaporedno številko skoka (vsak tekmovalec izvede več skokov).

2. naloga WikiLeaks

embassy (*id*, *name*)

privacy (*id*, *classification*)

cable (*id*, *date*, *header*, *content*, #*privacy_id*, #*embassy_id*)

reference (#*cable_id*, #*ref_cable_id*)

b)

Brez index	$O(n)$ MAX	$O(n)$ WHERE	$O(n \log n)$ ORDER BY
	95 ms	80 ms	80 ms
Index	$\sim 1\text{ms}$	$\sim 1\text{ms}$	$\sim 1\text{ms}$
	$O(\log(n))$	$O(\log(n))$	$O(1)$

(a) Ustvarite tabelo *wiki* ter ji določite primarni ključ

```
CREATE TABLE wiki AS SELECT c.id, date, name, classification FROM cable c JOIN embassy e ON  
c.embassy_id = e.id JOIN privacy p ON c.privacy_id = p.id WHERE date > 0;
```

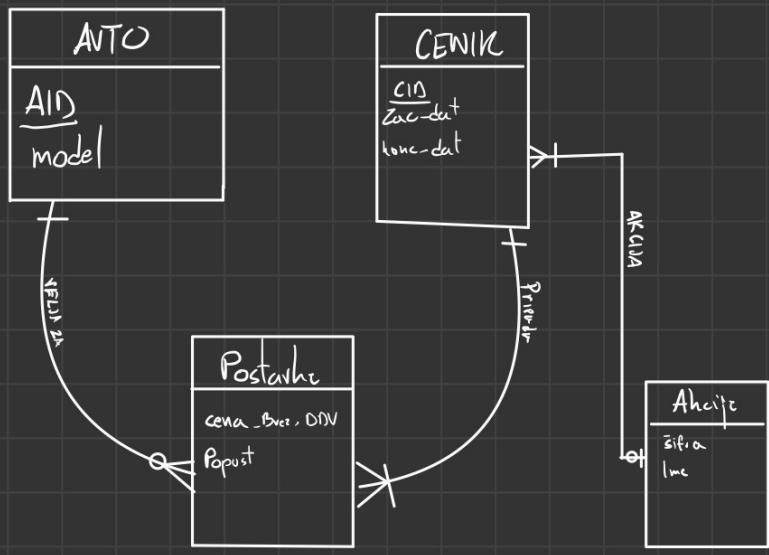
```
ALTER TABLE wiki ADD PRIMARY KEY(id)
```

(b) Ocenite čas izvajanja in zahtevnosti naslednjih poizvedb

```
SELECT [SQL_NO_CACHE] * FROM wiki WHERE date >= '2009-01-01' AND date <= '2009-12-31';
```

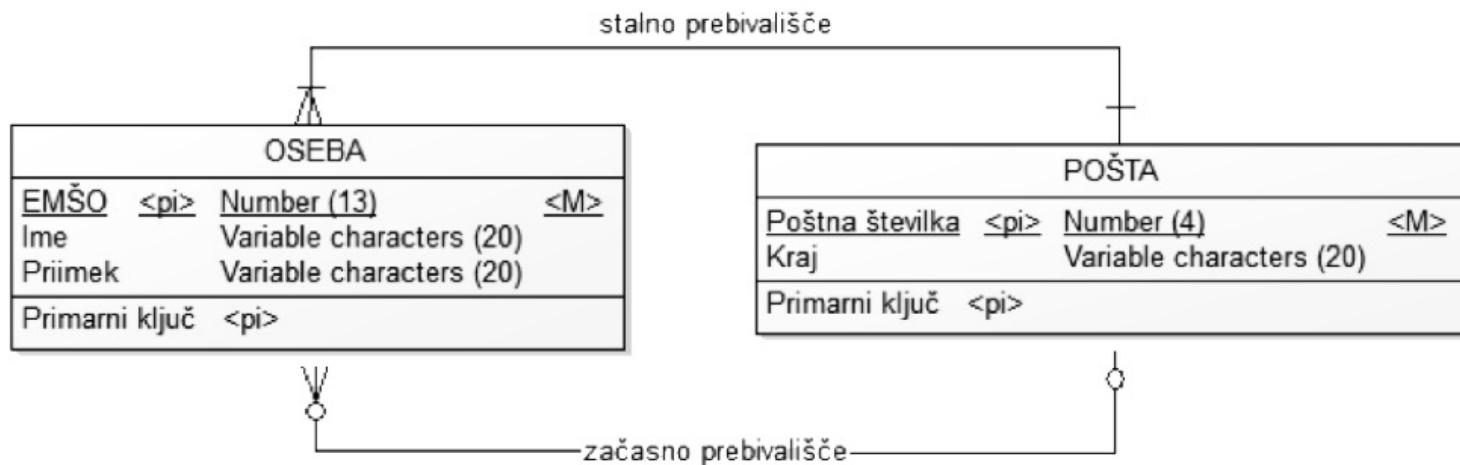
```
SELECT [SQL_NO_CACHE] * FROM wiki ORDER BY date;
```

```
SELECT [SQL_NO_CACHE] MAX(date) FROM wiki;
```



Koraki konceptualnega načrtovanja

1. identifikacija entitetnih tipov in povezav
 2. identifikacija atributov in določitev domen
 3. določitev entitetnih identifikatorjev tipov
 4. identifikacija šibkih entitetnih tipov
 5. uporaba elementov razširjenega modela
 6. preverjanje na podlagi zahtev in z uporabnikom



Konceptualni podatkovni model

Osnove podatkovnih baz

2. letnik univerzitetnega študija na FRI

2021/22

3. naloga WikiLeaks

embassy (id, name)

privacy (id, classification)

cable (id, date, header, content, #privacy_id, #embassy_id)

reference (#cable_id, #ref_cable_id)

*ALTER TABLE wiki
ADD INDEX (date)*

(a) Pohitrite izvajanje zgornjih poizvedb z indeksiranjem

(b) Ustvarite indeksa za atributa *name* in *classification*

(c) Ali je pohitritev s pomočjo indeksiranja brezplačna?

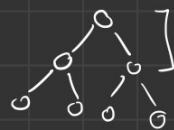
Ne, porabi bistveno več prostora

*Kupno časovno kompleksnost
v zmanj za prostorsko*

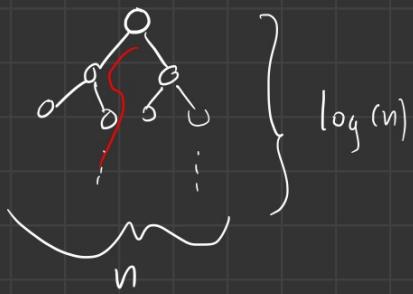
ID	ime	starost
1	A	10
2	B	30
3	C	15
4	D	20
5	E	42

starost = 20

10
|\
30
/ \
15 42
|
20

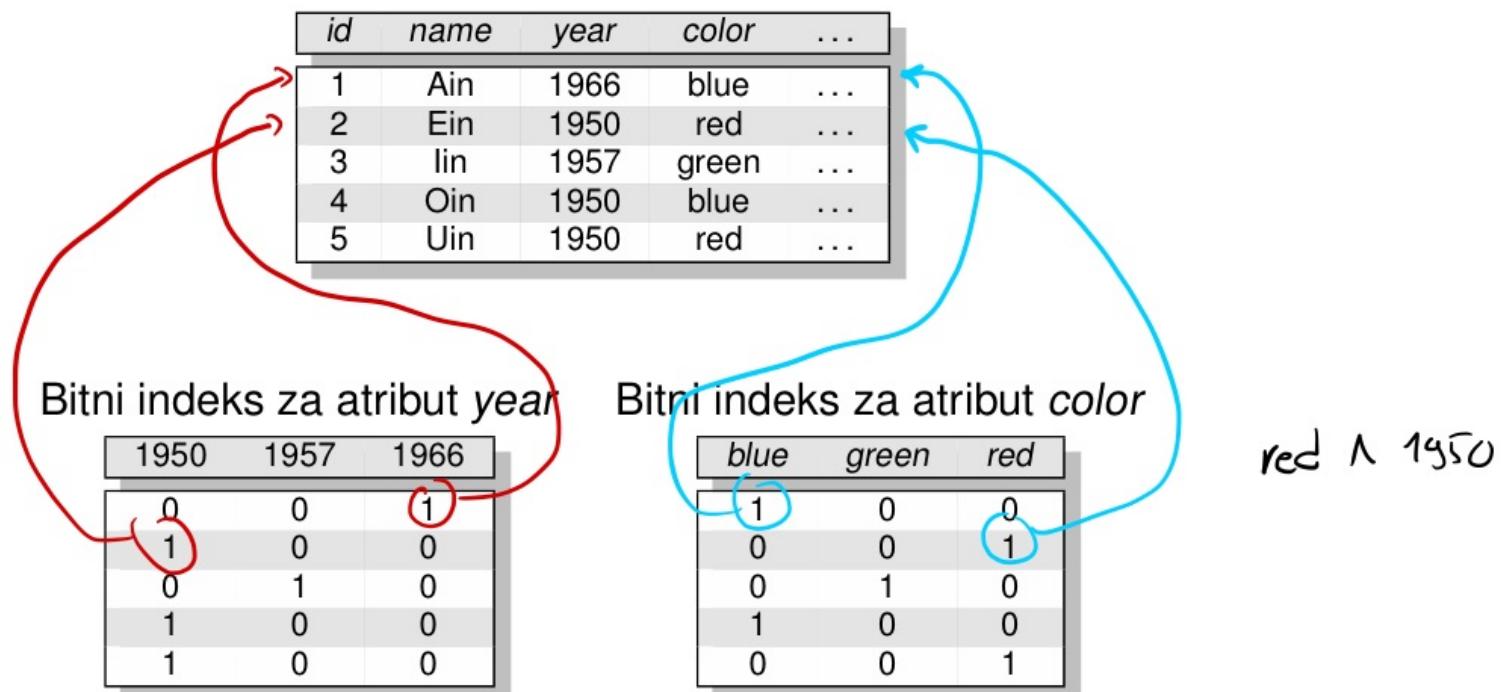


$d^h = \text{st. listov}$



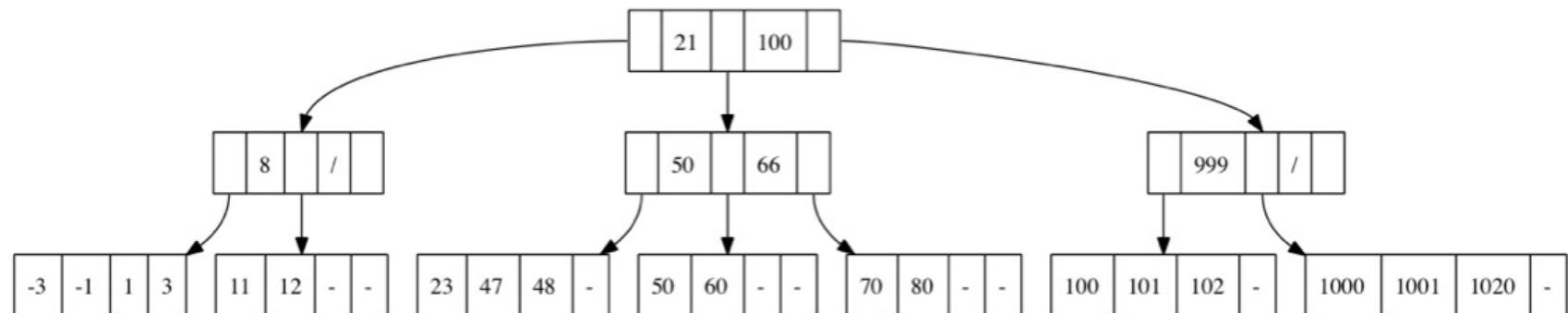
Bitni indeks

- ▶ bitni indeks je urejena binarna matrika
- ▶ celotna matrika se prilagaja pri posodabljanju zapisov
- ▶ časovna zahtevnost iskanja zapisov je *bitni* $O(n) < O(n)$



B⁺ indeks

- B^+ indeks je dinamično uravnoteženo drevo
 - celotno drevo se prilagaja pri posodabljanju zapisov
 - časovna zahtevnost osnovnih operacij $O(\log n) \ll O(n)$



B⁺ indeks za namišljeni atribut