Теоретски проект по Напредни бази на податоци

Високи нормални форми, 4-та, 5-та и домен-клуч

Димитриј Мијоски

Факултет за информатички науки и компјутерско инженерство, УКИМ, Скопје

Индекс 111132, имејл: [mijoski.dimitrij@students.finki.ukim.mk](mailto:mijoski.dimitrij@students.finki.ukim.mk)

**Резиме.** За овој проект ќе ги дефинираме и опишеме 4-та нормална форма, 5-та нормална форма и форма на домен-клуч. Заедно со нив и ќе ги дефинираме и опишеме на информативен и јасен начин повеќевредносни зависности, потребни за 4-та НФ, и зависности на спојување потребни за 5-та НФ.

**Клучни зборови:** нормализација, четврта нормална форма, петта нормална форма, нормална форма на домен-клуч, повеќевредносни зависности, зависности на спојување.

1. Вовед

Кога беше предложен релациониот модел во 1970 од Е. Ф. Код [1], во истиот труд беше предложен неформален процес на нормализација, а со тоа беше дефинирана и првата нормална форма. Таму е наведено дека една релација може да има сложен домен во кој елементите се исто така релација. Тоа значи дека елементот во некоја н-торка е множество од други н-торки. Било предложено од таквите релации сложените домени да се разделат како посебни релации и да се поврзат со првичната преку надворешни клучеви. Така би добиле табели (релации) каде сите домени се прости т.е. елементите во нивните н-торки се атомични, не се множество од вредности туку се една вредност. Ваквите нормални релации се многу полесни за преставување, можат да се претстават како дводимензионален низи односно табели (каде секоја ќелија содржи една вредност).

Тука да дадеме збор два за терминологијата [2]. Една релација се две работи, множество од н-торки (*инстанца*) и множество од *атрибути* (*шема*). Секоја н-торка има *елементи* т.е. *вредности на атрибути*. Сите можни вредности на еден атрибут го даваат *доменот* на тој атрибут. Па така зборовите атрибут и домен често се користат наизменично, иако формално се различни, атрибут е само име, идентификација или значење на вредностите во една н-торка, додека домен на атрибут се сите негови вредности.

Процесот на нормализација е математички формализиран заедно со функциските зависности [3]. Велиме дека кај ненормализирани табели или „до некаде“ нормализирани табели имаме повторување на податоци кое пак доведува до аномалии на внес, на ажурирање и на бришење и со нормализирање во повисока НФ отстрануваме некои аномалии. Потоа задаваме ограничување дека релационите бази на податоци мора да содржат табели во прва нормална форма, односно не дозволуваме сложени домени т.е. повеќевредносни атрибути. Тука пак имаме повторување на податоци (само сме овозможиле релацијата да се да се прикажува во правоаголна табела) и пак може да се јават аномалиите. Над табели во прва нормална форма дефинираме функциска зависност која многу јасно ни кажува за кои атрибути може да видиме повторување. Па секоја нормална форма е дефинирана преку постоењето на само одредени типови на ФЗ врз табела и непостоењето на сите останати. За проектов претпоставуваме дека читателот е запознаен со прва, втора трета и нормална форма БКНФ, со нивните дефиниции преку функциски зависности или барем нивните информативни и неформални дефиниции, и со теоријата на ФЗ која е понатаму развиена.

Од теорија на ФЗ посебно ќе ги спомнам 3-те аксиоми на Армстронг, 3-те изведени својства од нив, затворач или обвивка на ФЗ и минимална т.е. канонична покривка.

Во понатамошниот тек ќе дефинираме повеќевредносни зависности кои ни се потребни пак за дефиниција на 4-та нормална форма и зависности па проекција и спојување кои ни се потребни за 5-та нормална форма. На самиот крај ќе зборуваме и за НФ домен-клуч.

1. Факти со една вредност и факти со повеќе вредности
2. 4-та нормална форма
3. Референци
4. Codd, Edgar Frank. “A relational model of data for large shared data banks.” Communications of the ACM 13.6 (1970): 377-387.
5. Yang, Chao-Chih. “Relational Databases.” Englewood Cliffs, New Jersey: Prentice-Hall, 1986. 12-19.
6. Codd, Edgar Frank. “Further Normalization of the Data Base Relational Model.”

# Референци

Codd, Edgar Frank. “A relational model of data for large shared data banks.” *Communications of the ACM 13.6* (1970): 377-387.

Yang, Chao-Chih. “Relational Databases.” Englewood Cliffs, New Jersey: Prentice-Hall, 1986. 12-19.