

9. Нулеви стойности

Лекционен курс "Бази от данни"

Определение

До сега обяснявахме базовите идеи, свързани с КК и ВК в релационния модел, но съществува един фактор, който беше досега игнориран – нулевите стойности:

Нулева стойност: липса на стойност (данни).

- В реалния свят има много такива ситуации - напр. неизвестен адрес;
- В релационния модел е приет един специален маркер за означаване на липсваща информация – NULL;
- Един атрибут може да съдържа или може да му бъде забранено да съдържа нулеви стойности.

Кандидат-ключове и нулеви стойности

Правило за цялостност на обект : не се разрешава на нито един от компонентите на ПК да приема нулеви стойности.

Причини:

- базова релация - (кореспондират с понятието таблица в SQL) автономна именувана релация. Базовите релации кореспондират с обекти от реалния свят, които трябва да бъдат идентифицирани по някакъв начин;
- първичните ключове изпълняват функцията на уникалната идентификация в релационния модел, следователно един първичен ключ, който е напълно нулев ще бъде едно противоречие в този контекст.

Обобщение

В една релационна БД ние не трябва да записваме информация, която не можем да идентифицираме.

Забележки

- Правилото се прилага само за базови релации - другите релации могат да имат кандидат-ключове, на които е разрешено NULL;
- Правилото се прилага само за ПК - АК (алтернативен ключ) могат да имат NULL. Един АК, на който е разрешено, не може да бъде избран за ПК.

Външни ключове и нулеви стойности

Да разгледаме примера “Държава-Регион”.

Възможно е някоя държава да не е причислена към конкретен географски регион \Rightarrow допуска се ВК да има нулеви стойности.

Забележка: за всеки ВК конструкторът на БД трябва да реши дали се допускат нулеви стойности.

Дефиниция на ВК

Ако външният ключ позволява нулеви стойности, то дефиницията му може да бъде малко променена.

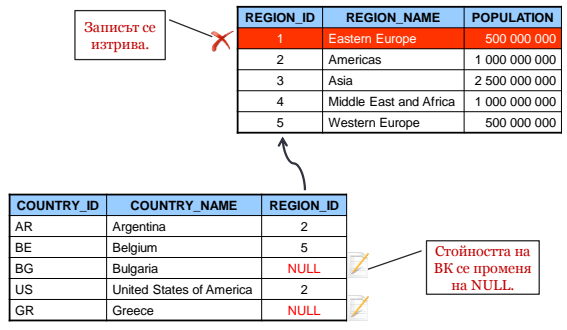
Външен ключ: нека R е базова релация. Външен ключ (ВК) в R е едно подмножество на множеството на атрибутите $ВК = \{A_1^R, A_2^R, \dots, A_n^R\}$ така, че:

- съществува една базова релация R_1 с ключ-кандидат;
- във всеки един момент стойностите на ВК в R са **нулеви** или равни на стойности на КК в някой запис на R_1 .

Да се върнем на въпроса какво ще стане, когато искаме да изтрием п-торка от целевата релация, например, да изтрием регион, за който съществува държава:

- Ако за ВК са разрешени NULL - съществува нова трета възможност – NULLIFIES - т.е. ВК получава стойност NULL и тогава се изтрива основната п-торка;
- Същото важи и за промяната.

NULLIFIES



NULL и логическите оператори

- SQL има три логически стойности: TRUE, FALSE и UNKNOWN;
- Стойността UNKNOWN произлиза в резултат използването на NULL в сравнения и с други предикати, но UNKNOWN е логическа стойност и не е еквивалентна на NULL, което е маркер за липса на стойност;
- Ето защо в SQL пишем
 - ✗ IS [NOT] NULL
 - вместо
 - ✗ = NULL

NULL и операторите

| AND | TRUE | UNKNOWN | FALSE |
|---------|---------|---------|-------|
| TRUE | TRUE | UNKNOWN | FALSE |
| UNKNOWN | UNKNOWN | UNKNOWN | FALSE |
| FALSE | FALSE | FALSE | FALSE |

| OR | TRUE | UNKNOWN | FALSE |
|---------|------|---------|---------|
| TRUE | TRUE | TRUE | TRUE |
| UNKNOWN | TRUE | UNKNOWN | UNKNOWN |
| FALSE | TRUE | UNKNOWN | FALSE |

| x | NOT |
|---------|---------|
| TRUE | FALSE |
| UNKNOWN | UNKNOWN |
| FALSE | TRUE |

| | |
|-------------|-----------|
| x * UNKNOWN | = UNKNOWN |
| x / UNKNOWN | = UNKNOWN |
| x + UNKNOWN | = UNKNOWN |
| x - UNKNOWN | = UNKNOWN |
| x % UNKNOWN | = UNKNOWN |

Релационен интегритет - обобщение

Ограниченията, свързани с интегритета на данните, се отнасят към коректността на данните в базата данни. Интегритетът на данните предоставя механизъм за поддържане на консистентността на данните за операциите INSERT, UPDATE и DELETE;

Друг тип ограничения за интегритет на данните са:

- Цялостност на обект (Entity integrity);
- Поддръжка на нулеви стойности (NULL integrity);
- Ограничения, свързани с областите (Domain integrity);
- Референциален интегритет (Referential integrity).

Цялостност на обект (Entity integrity)

Цялостност на обект означава, че идентификаторът (ПК) не може да приема нулеви стойности.

- Първичният ключ в релация уникално идентифицира ред. Цялостността на обект означава, че за да бъде представен обект в базата данни той трябва да бъде напълно идентифицируем чрез атрибутите му, съставлящи първичния ключ;
- Да вземем за пример релацията STUDENT с атрибути FacNo, Name, Year. Нека FacNo да е ПК. Не можем да въведем данни за студент без да въведем неговия факултетен номер, защото само той го идентифицира напълно, което значи, че ПК не може да има липсваща стойност.

Поддръжка на нулеви стойности

Нулева стойност означава неизвестни данни.

- Релацията STUDENTS има атрибут Name, който не може да бъде неизвестен (NOT NULL) – в реалността няма записан студент без име.

Ограничения, свързани с областите

- Областите в релационния модел се използват за дефиниране на характеристиките на колоните в таблиците;
- Те представят множества от допустими стойности за атрибутите;
- Областите се специфицират с име, тип данни, върху който са дефинирани и диапазон в този тип данни;
- Ограниченията, свързани с областите, се представят с дефиниране на списък от позволени стойности, с диапазон или с израз, който проверява дали стойността е допустима;
- Стриктно казано, само стойности от една и съща област могат да бъдат сравнявани или обединявани с оператор UNION.

Примери:

- ✓ възрастта на човек не може да съдържа букви или да е отрицателно число;
- ✓ полът на човек може да се представи със стойностите (М, Ж), като всяка друга стойност би нарушила ограничението

Референциален интегритет

- В релационния модел връзките между таблиците са дефинирани с използването на външни ключове;
- Правилото за референциалния интегритет гласи, че базата данни не трябва да съдържа стойности във ВК, които нямат съответни в референцираните ПК;
- То също така не предполага, че ВК не може да бъде NULL – това решение се оставя за дизайнера, който преценява в зависимост от логиката;
- Т.е. правилото гласи, че стойностите на всеки ВК трябва да са измежду стойностите на ПК от друга (не е задължително) таблица или да бъдат NULL.