УДК 519.72
МРНТИ 27.03.66, 27.47.19

https://doi.org/10.55452/1998-6688-2022-19-2-6-12

О ЗАПРОСАХ БАЗ ДАННЫХ НАД СЛАБО О-МИНИМАЛЬНОЙ ОБЛАСТЬЮ С МАЛЫМ СЧЕТНЫМ СПЕКТРОМ

АЛТАЕВА А.Б.

Казахский национальный университет имени аль-Фараби,
050040, г.Алматы, Казахстан

Аннотация. *В реляционной модели баз данных, введенной Э.Ф. Коддом, состояние базы данных понимается как конечная совокупность отношений между элементами. Имена отношений и их арности (местности) фиксируются и называются схемой базы данных. Отдельная информация, хранимая в отношениях данной схемы, называется состоянием базы данных. Хотя реляционные базы данных были придуманы для конечных совокупностей данных, часто удобно предполагать, что существует бесконечная область определения.*
Мы исследуем реляционные базы данных над упорядоченной областью определения с некоторыми дополнительными отношениями – типичным примером является множество рациональных чисел с отношением линейного порядка и бинарной операцией сложения. Если в качестве языка запросов используется язык логики предикатов первого порядка, то запросы могут использовать как отношения базы данных, так и отношения области определения, при этом переменные изменяются на всей области определения. В фокусе наших исследований запросы первого порядка (FO), инвариантные относительно перестановок, сохраняющих порядок, такие запросы называются порядково-генерическими. Установлено, что для некоторых областей порядково-генерические запросы первого порядка сводятся к запросам чистого порядка. Здесь мы доказываем теорему сводимости над слабо о-минимальной областью определения, имеющей ранг выпуклости 1 и малый счетный спектр.

Ключевые слова: *слабая о-минимальность, запрос баз данных, счетный спектр, почти омега-категоричность, ранг выпуклости.*

КІШІ ЕСЕПТІК СПЕКТРІ БАР ӘЛСІЗ О-МИНИМАЛДЫ АНЫҚТАУ ОБЛЫСЫ БОЙЫНША МӘЛІМЕТ БАЗАСЫ СҰРАУЛАР ТУРАЛЫ

АЛТАЕВА А.Б.

әл-Фараби атындағы Қазақ Ұлттық университеті
050040, Алматы қ, Қазақстан

Аңдатпа. *Реляциялық деректер қоры моделінде Э. Ф. Кодд, дерекқор күйі элементтер арасындағы қарым-қатынастардың соңғы жиынтығы ретінде түсініледі. Қатынастардың* *атаулары және олардың аритеттері (орындары) бекітілген және деректер қоры схемасы деп аталады. Бұл схеманың қатынастарында сақталатын жеке ақпарат мәліметтер қорының күйі деп аталады. Реляциялық дерекқорлар деректердің ақырғы жиындарына арналған болса да, көбінесе шексіз анықтау облысы бар деп болжауға ыңғайлы. Біз реляциялық мәліметтер базасын қосымша қатынастармен реттелген анықтау облысы бойынша зерттейміз – типикалық мысалы қосу бинарлық операциямен рационалдық сандар реттелген жиыны болады. Бірінші ретті (FO) предикат логикалық тіл сұрау тілі ретінде пайдаланылса, сұраулар барлық доменде өзгеретін айнымалы мәндермен дерекқор қатынастарын да, домен қатынастарын да пайдалана алады. Біздің зерттеуіміз бірінші ретті сұраныстарға бағытталған, олар ретті сақтайтын ауыстыруға қатысты инвариантты болады, мұндай сұраныстар рет-генерикалық деп аталады. Кейбір облыстар үшін бірінші ретті рет-генерикалық сұраныстар таза реттік сұраныстарға дейін азайтылатыны анықталды. Мұнда біз кіші есептік спектрі бар және дөңестік рангісі 1 болған әлсіз о-минималды анықтау облысы бойынша редукция теоремасын дәлелдейміз.*

Түйін сөздер: *әлсіз о-минималдық, мәлімет базасының сұранысы, есептік спектрі, омега-категориялық дерлік, дөңестік рангісі.*

ON DATABASE QUERIES OVER A WEAKLY O-MINIMAL DOMAIN WITH A SMALL COUNTABLE SPECTRUM

ALTAYEVA A.B.

Al-Farabi Kazakh National University,
050040, Almaty, Kazakhstan

Abstract. *In the relational database model introduced by E.F. Codd, the database state is understood as a finite set of relationships between elements. The names of the relationships and their arnts (locations) are fixed and called the database schema. The individual information stored in the relationships of this schema is called the state of the database. Although relational databases were devised for finite sets of data, it is often convenient to assume that there is an infinite domain. We consider relational databases organized over an ordered domain with some additional relations – a typical example is the set of rational numbers together with the relation of linear order and binary operation of addition. If the first-order predicate logic language is used as the query language, then queries can use both database relationships and domain relationships, with variables changing throughout the domain. In the focus of our study are the first-order (FO) queries that are invariant under order-preserving permutations – such queries are called order-generic. It was discovered that for some domains order-generic FO queries fail to express more than pure order queries. Here we prove the collapse result theorem over a weakly o-minimal domain having convexity rank 1 and a small countable spectrum.*

Keywords: *weak o-minimality, database query, countable spectrum, almost omega-categoricity, convexity rank.*

Введение

В реляционной модели баз данных, введенной Э.Ф. Коддом [1–2], состояние базы данных понимается как конечная совокупность отношений между элементами. Имена отношений и их арности (местности) фиксируются и называются схемой базы данных. Отдельная информация, хранимая в отношениях данной схемы, называется состоянием базы данных. Хотя реляционные базы данных были придуманы для конечных совокупностей данных, часто удобно предполагать, что существует бесконечная область определения – например, целые или рациональные числа, – так что элементы данных выбираются из этой области. Функции и отношения, определенные на всей области определения (например, отношение линейного порядка < и операция сложения +), могут быть также использованы при запрашивании. Например, если в качестве языка запросов используется язык логики предикатов первого порядка, то запросы могут использовать как отношения базы данных, так и отношения области определения, при этом переменные изменяются на всей области определения. Выразительная сила запросов баз данных исследовалась в работах [3–10].

Формальная постановка

Пусть M – бесконечная структура сигнатуры L. Здесь мы рассматриваем упорядоченные структуры. Это означает, что L включает бинарный реляционный символ <, интерпретация которого в M удовлетворяет аксиомам линейного порядка. Мы фиксируем схему базы данных SC и вводим следующие обозначения:

$$L_0 = \{<\},\ L` = L_0 \cup SC,\ L`` = L \cup SC.$$

Запрос базы данных может быть формально определен как отображение, которое принимает состояние базы данных и производит новое отношение фиксированной арности над M. Мы рассматриваем два языка для запрашивания. Запросы первого языка есть формулы сигнатуры L` – мы называем их ограниченными. Запросы второго языка есть формулы сигнатуры L`` – мы называем их расширенными.

Итак, базы данных предназначены для хранения текущей информации о как-то структурированной предметной области. В каждый момент времени эта информация является конечной и представляет собой конечный набор конечных таблиц. Обычно число таблиц и устройство каждой таблицы не меняются с течением времени, но меняются строки таблиц. Могут добавляться новые строки и удаляться некоторые старые. Строки хранящихся таблиц представляют собой конечные последовательности элементов. Число элементов каждой последовательности фиксировано для фиксированной таблицы. Устройство таблицы практически и есть число элементов в каждой строке этой таблицы. Более формально, каждая таблица – это конечно местное конечное отношение, а сама база данных – это конечный набор конечно местных конечных отношений. Для удобства разговора о базе данных каждому ее отношению приписывают некоторое имя с указанием числа аргументов (или местности) этого имени отношения. Схема (или сигнатура) базы данных и есть конечная последовательность этих имен отношений с указанием местности каждого имени. В каждый момент времени именам отношений из этой схемы присвоены некоторые отношения соответствующих местностей. Это – состояние базы данных в данный момент.

Состояние называется конечным, если все его отношения конечны. Иногда удобно рассматривать не произвольные состояния базы данных, а ограниченные какими-то условиями. Типичным ограничением является условие, что элементы всех строк всех таблиц выбраны из фиксированного подмножества I универсума. Другими словами, каждому имени отношения из рассматриваемой схемы базы данных поставлено в соответствие отношение той же местности на множестве I. В этом случае говорят, что рассматриваемое состояние базы данных является состоянием над I.

Мы будем рассматривать локально генерические запросы, которые являются инвариантами при любых сохраняющих линейное упорядочение отображениях конечных подмножеств универсума в универсум. Грубо говоря, ответ на такой запрос основывается на хранящейся информации, но не зависит от способа кодировки этой информации при хранении.

Определение 1. Будем говорить, что k–арный запрос Θ является локально генерическим над конечными состояниями, если $\bar{a} \in \Theta$ тогда и только тогда, когда $\varphi(\bar{a}) \in \Theta(\varphi(s))$ для любого частичного <-изоморфизма $\varphi : X \to M$, где $X \subseteq M$ для любого конечного состояния s над X и для любого k-кортежа $\bar{a}$ в X.

Состояние s обогащает универсум *M* сигнатуры *L* до *L''*-структуры, которую мы будем обозначать как *(M, s)*.

Определение 2. ρ-состояние s для *L*-структуры *W* называется псевдоконечным в *W*, если *(W,s)* есть модель *L''*-теории первого порядка всех структур *(W, r)*, где *r* – конечное состояние над *W*.

Псевдоконечное множество – это частный случай псевдоконечного состояния. Имеется в виду сигнатура, состоящая из одного одноместного отношения и некоторых других отношений. Рассматриваются такие системы этой сигнатуры, на которых выполняются все замкнутые формулы логики предикатов, истинные на всех конечных системах этой сигнатуры. Тогда интерпретация этого одноместного отношения в такой системе называется псевдоконечным множеством.

Определение 3. Будем говорить, что полная теория T имеет Свойство Изоляции, если существует кардинал λ такой, что для любого псевдоконечного множества A и для любого элемента $\bar{a}$ модели теории T существует $A_0 \subseteq A$ такое, что $|A_0| < \lambda$ и $tp(\alpha/A_0)$ изолирует $tp(\alpha/A)$.

Для произвольных подмножеств A, B структуры M пишут A < B, если a < b всякий раз, когда a ∈ A и b ∈ B. Если A ⊂ M и x ∈ M, то пишут A < x, если A < {x}. Для произвольного полного 1-типа p через p(M) обозначают множество реализаций типа p в M. Открытым интервалом I в структуре M называется параметрически определимое подмножество структуры M вида I = {c ∈ M: M | = a < c < b} для некоторых a, b ∈ M ∪ {– ∞, ∞}, где a < b. Аналогично можно определить замкнутые, полуоткрытые-полузамкнутые и т.п. интервалы в M, так что, например, произвольная точка структуры M является сама (тривиальным) замкнутым интервалом. Подмножество A структуры M называется выпуклым, если для любых a, b ∈ A и c ∈ всякий раз, когда a < c < b следует, что c ∈ A.

Данная статья касается понятия слабой о-минимальности, первоначально глубоко исследованного Д. Макферсоном, Д. Маркером и Ч. Стайнхорном в [11]. Слабо о-минимальная структура есть линейно упорядоченная структура $M = \langle M, =, <, \ldots \rangle$, такая, что любое определимое (с параметрами) подмножество структуры M является объединением конечного числа выпуклых множеств в M. Вспомним, что такая структура M называется о-минимальной, если каждое определимое (с параметрами) подмножество структуры M является объединением конечного числа интервалов в M. Таким образом, слабая о-минимальность является обобщением о-минимальности. Ранг выпуклости формулы с одной свободной переменной введен в [12]. Ниже мы расширяем определение ранга выпуклости формулы [12] на произвольные множества (необязательно определимые):

Определение 4. [12] Пусть T – слабо о-минимальная теория, M – достаточно насыщенная модель теории T, A ⊆ M. Ранг выпуклости множества A (RC(A)) определяется следующим образом:

1) RC(A) = –1, если A = ∅.

2) RC(A) = 0, если A конечно и непусто.

3) RC(A) ≥ 1, если A бесконечно.

4) RC(A) ≥ α + 1, если существует параметрически определимое отношение эквивалентности E(x, y) такое, что существуют bi ∈ A, i ∈ ω, которые удовлетворяют следующему:

• для любых i, j ∈ ω, всякий раз когда i ≠ j, мы имеем M ⊨ ¬E(bi, bj)

• для каждого i ∈ ω RC(E(M, bi)) ≥ α и E(M, bi) – выпуклое подмножество множества A

5) RC(A) ≥ δ, если RC(A) ≥ α для всех α ≤ δ (δ предельный).

Если RC(A) = α для некоторого α, то мы говорим, что RC(A) определяется. В противном случае (т.е. если RC(A)) ≥ α для всех α), мы полагаем, RC(A) = ∞.

Ранг выпуклости формулы ϕ(x, a ̅), где

$a \in M$, определяется как ранг выпуклости множества $\phi(M, \bar{a})$, т.е. $RC(\phi(x, \bar{a})) := RC(\phi(M, \bar{a}))$. Ранг выпуклости 1-типа p определяется как ранг выпуклости множества p(M), т.е. RC(p):= RC(p(M)).

В частности, теория имеет ранг выпуклости 1, если не существует определимого (с параметрами) отношения эквивалентности с бесконечным числом выпуклых бесконечных классов. Очевидно, что о-минимальная теория имеет ранг выпуклости 1.

Определение 5. [13] Пусть M – слабо о-минимальная структура, $A, B \in M$, M – $|A|^+$-насыщенна, $p, q \in S_1(A)$ – неалгебраические. Будем говорить, что тип p не является слабо ортогональным типу q, если существуют A-определимая формула H(x, y), $\alpha \in p(M)$ и $\beta_1, \beta_2 \in q(M)$ такие что $\beta_1 \in H(M, a)$ и $\beta_2 \notin H(M, \alpha)$.

Определение 6. [14, 15] Пусть T – полная теория, $p_1(x_1), \dots, p_n(x_n) \in S_1(\emptyset)$. Тип $q(x_1, \dots, x_n) \in S_n(\emptyset)$ называется $(p_1, \dots, p_n)$-типом, если $q(x_1, \dots, x_n) \supseteq \cup_{i=1}^{n} p_i(x_i)$. Множество всех $(p_1, \dots, p_n)$-типов теории T обозначается через $S_{p_1,\dots,p_n}(T)$. Счетная теория T называется почти омега-категоричной, если для любых типов $p_1(x_1), \dots, p_n(x_n) \in S_1(\emptyset)$ существует лишь конечное число типов $q(x_1, \dots, x_n) \in S_{p_1,\dots,p_n}(T)$.

Почти омега-категоричность тесно связана с понятием эренфойхтовости теории. Так, в работе [14] доказано, что если T – почти омега-категоричная теория, имеющая ровно три счетные попарно неизоморфные модели, то в теории T интерпретируется плотный линейный порядок. Тем не менее существует пример (построенный М.Г. Перетятькиным в [16]) теории, имеющей ровно три счетные попарно неизоморфные модели, но не являющейся почти омега-категоричной.

В работе [17] установлены почти омега-категоричность эренфойхтовых вполне о-минимальных теорий и выполнимость принципа замены для алгебраического замыкания для почти омега-категоричных вполне о-минимальных теорий. Недавно были доказаны ортогональность любого семейства попарно слабо ортогональных неалгебраических 1-типов над пустым множеством для таких теорий и бинарность почти омега-категоричных вполне о-минимальных теорий [18] и почти омега-категоричных слабо о-минимальных теорий ранга выпуклости 1 [19].

В настоящей работе исследуется проблема выразимости расширенных запросов через ограниченные над слабо о-минимальной областью определения баз данных, имеющей ранг выпуклости 1 и малый счетный спектр. Мы доказываем, что слабо о-минимальная теория ранга выпуклости 1 с малым счетным спектром имеет свойство изоляции. В качестве следствия мы получаем сводимость расширенных запросов к ограниченным над слабо о-минимальной областью определения, имеющей ранг выпуклости 1 и малый счетный спектр.

Результаты

Теорема 7. [3] Предположим, что теория первого порядка структуры M имеет Свойство Изоляции. Пусть расширенный запрос φ эквивалентен над конечными состояниями ограниченному запросу.

Теорема 8. Пусть T – слабо о-минимальная теория ранга выпуклости 1 с малым счетным спектром. Тогда T имеет Свойство Изоляции.

Доказательство теоремы 9. Заметим, что слабо о-минимальная теория ранга выпуклости 1 с малым счетным спектром является почти омега-категоричной. Пусть M – достаточно насыщенная модель теории T. Возьмем произвольные элемент $a \in M$ и бесконечное множество $A \subseteq M$ и рассмотрим $p(x) := tp(a / A)$. В силу слабой о-минимальности p(M) выпукло, и, следовательно, тип p(x) определяется выпуклыми формулами.

Случай 1. p(x) – изолированный. Тогда существует формула $\varphi(x, \bar{b})$, где $\bar{b} \in A$, такая, что $\varphi(M, \bar{b})$ выпукло и $p(M) = \varphi(M, \bar{b})$. Таким образом, в качестве A_0 можем взять множество элементов из кортежа $\bar{b}$.

Случай 2. p(x) – квазирациональный. Не умаляя общности, предположим, что p(x) – квазирациональный вправо. Тогда существует выпуклая формула $U(x, \bar{b})$ для некоторого $\bar{b} \in A$, так что $p(M) \subseteq U(M, \bar{b})$ и

$U(M,\bar{b})^+ = p(M)^+$. В силу бинарности T для любой выпуклой формулы $\varphi_i(x,\bar{b}_i) \in p$ левая граница множества $\varphi_i(M,\bar{b}_i)$ определяется выпуклой формулой $\varphi_i^1(x,b_i^1)$ для некоторого $b_i^1 \in \bar{b}_i$. В силу почти омега-категоричности попарно неэквивалентных выпуклых формул $\theta(x,b_i^1)$ с условием $p(M) \subseteq \theta(M,b_i^1)$ конечное число. Таким образом, мы заключаем, что левая граница множества p(M) определяется счетным числом констант из A. Поэтому в качестве A_0 можем взять счетное подмножество множества A.

Случай 3. p(x) – иррациональный. В этом случае можно показать аналогично случаю 2, что как левая, так и правая границы множества p(M), определяются счетным λ множеством констант из A.

Таким образом, в качестве λ можем взять первый несчетный кардинал ω_1. Следовательно, T имеет Свойство Изоляции.

Заключение

Таким образом, в качестве следствия получаем, что если T – слабо о-минимальная теория ранга выпуклости 1 с малым счетным спектром, то любой расширенный запрос, являющийся локально генерическим над конечными состояниями, эквивалентен ограниченному запросу.

Благодарности

Данные исследования поддержаны Комитетом науки Министерства образования и науки Республики Казахстан (грант AP08855544).

REFERENCES

1 Codd E.F. A relational model for large shared data banks // Communications ACM, 1970, vol. 13, no. 6, pp. 377–387.

2 Codd E.F. Relational completeness of database sublanguages // Database systems, Prentice-Hall, 1972, pp. 33–64.

3 Benedikt M., Dong G., Libkin L., Wong L. Relational expressive power of constraint query languages // Journal of ACM, 1998, vol. 45, no. 1, pp. 1–34.

4 Belegradek O.V., Stolboushkin A.P. and Taitslin M.A. Extended order-generic queries // Annals of Pure and Applied Logic, 1999, vol. 97, pp. 85–125.

5 Taitslin M.A. A general condition for collapse results // Annals of Pure and Applied Logic, 2002, vol. 113, no. 1–3, pp. 323–330.

6 Dudakov S.M., Tajclin M.A. Transljacionnye rezul'taty dlja jazykov zaprosov v teorii baz dannyh // Uspehi matematicheskih nauk, 2006, vol. 61, no. 2, pp. 3–66.

7 Kulpeshov B.Sh. On Problem of Expressiveness of Database Queries // International Journal of Mathematics, Computer Sciences and Information Technology, 2010, vol. 3, no. 2, pp. 123–128.

8 Kulpeshov B.Sh. To Reducibility of Database Queries over an Ordered Domain // Computer Modelling and New Technologies, 2012, vol. 16, no. 2, pp. 34–39.

9 Kulpeshov B.Sh. On Reducibility of database queries over a circularly minimal domain // Advances in Computational Sciences and Technology, 2013, vol. 6, no. 1, pp. 25–33.

10 Baizhanov B.S., Kulpeshov B.Sh. On the Isolation Property over a Database Domain // Journal of Mathematics and System Science, 2013, vol. 3, no. 2, pp. 96–100.

11 Macpherson H.D., Marker D., Steinhorn Ch. Weakly o-minimal structures and real closed fields // Transactions of the American Mathematical Society, 2000, vol. 352, pp. 5435–5483.

12 Kulpeshov B.Sh. Weakly o-minimal structures and some of their properties // The Journal of Symbolic Logic, 1998, vol. 63, pp. 1511–1528.

13 Baizhanov B.S. Expansion of a model of a weakly o-minimal theory by a family of unary predicates // The Journal of Symbolic Logic, 2001, vol. 66, pp. 1382–1414.

14 Ikeda K., Pillay A., Tsuboi A. On theories having three countable models // Mathematical Logic Quarterly, 1998, vol. 44, no. 2, pp. 161–166.

15 Sudoplatov S.V. Classification of countable models of complete theories. Part 1. Novosibirsk: Novosibirsk State Technical University Publishing House, 2018, 326 p. ISBN 978-5-7782-3527-4.

16 Peretyat'kin M.G. A theory with three countable models // Algebra and Logic, 1980, vol. 19, no. 2, pp. 139–147.

17 Kulpeshov B.Sh., Sudoplatov S.V. Linearly ordered theories which are nearly countably categorical // Mathematical Notes, 2017, vol. 101, no. 3, pp. 475–483.

18 Altayeva A.B., Kulpeshov B.Sh. Binarity of almost ω-categorical quite o-minimal theories // Siberian Mathematical Journal, 2020, vol. 61,no. 3, pp. 379–390.

19 Kulpeshov B.Sh., Mustafin T.S. Almost ω-categorical weakly o-minimal theories of convexity rank 1 // Siberian Mathematical Journal, 2021, vol. 62, no. 1, pp. 52–65.

Сведения об авторе

Алтаева Айжан Бакаткалиевна

Магистр, докторант Казахского национального университета имени аль-Фараби, пр. Аль-Фараби, 71, 050040, г. Алматы, Казахстан

ORCID ID: 0000-0001-9238-7131

E-mail: vip.altayeva@mail.ru

Авторлар туралы мәлімет

Алтаева Айжан Бакаткалиқызы, магистр, әл-Фараби атындағы Қазақ Ұлттық университетінің PhD докторанты, әл-Фараби даңғылы, 71, 050040, Алматы қ., Қазақстан

ORCID ID: 0000-0001-9238-7131

E-mail: vip.altayeva@mail.ru

Information about author

Altayeva Aizhan Bakatkalievna

Master, PhD student of Al-Farabi Kazakh National University, 71, al-Farabi avenu, 050040, Almaty, Kazakhstan

ORCID ID: 0000-0001-9238-7131

E-mail: vip.altayeva@mail.ru