

НОРМАЛИЗАЦИЯ АДРЕСА

стр. 46 – 50

С. Л. Гладков

ООО «Айгео»:

660079, Россия, г. Красноярск, ул. 60-летия Октября, д. 105

Контактные данные: С. Л. Гладков, e-mail: gladkovs@list.ru

Резюме: Цель. Ввести конструктивное понятие нормализации адреса. Это вторая статья из задуманной серии, посвященной проблеме использования адреса в качестве идентификатора местоположения объекта недвижимости. В статье анализируются нормализация адресов и исследуются шаги приведения произвольного адреса к нормальной форме. Материалы. Исходными материалами для статьи служат федеральные и региональные нормативные акты. Методы. Анализ и логика. Результаты выводы (заключение): Сформулировано формальное определение множества нормальных адресов. Предложены 4-х шаговые алгоритмы нормализации значений реквизитов и адресов.

Ключевые слова: нормализованный адрес, нормализация адреса, адрес, модель адреса, семантика адреса, реквизиты адреса, уникальность адреса, концептуальная модель

ADDRESS NORMALIZATION

pages 46 – 50

S. L. Gladkov

LLC Aigeo: 105, ul. 60th anniversary of October, Krasnoyarsk, 660079, Russia

Summary: Purpose. To introduce a constructive concept normalization of addresses. This is the second article of the series, devoted to the problem of using the address as an identifier of the location of the object of immovable into database. The article analyzes the normalization of addresses and studied steps of transformation address into its normal form. Materials. The source materials for the article are Federal and regional normative acts. Methods. Analysis and logic. The results, the findings (conclusion). The formal definition of the set of normal addresses introduced. 4-step algorithms for normalization of requisites and addresses are proposed.

Keywords: the normalized address, the normalization of addresses, address, address model, address semantics, address requisites, address uniqueness, conceptual model

Нормализация адреса

1, Множества нормальных адресов и реквизитов

Множеством **нормальных адресов** будем называть объединение базовых и аннулированных адресов

$$\bar{A} \equiv A_b \cup A_c, \quad (1)$$

где $A_b \equiv A$ – множество базовых адресов, обладающих свойствами уникальности, обязательности легитимности, актуальности и однозначности, A_c – множество аннулированных (отмененных) адресов. Множества базовых и аннулированных адресов не пересекаются $A_b \cap A_c = \emptyset$. F – функция взаимно-однозначно отображает множество базовых адресов на множество объектов адресации:

$$A_b \xleftrightarrow{F} O_b \quad (2)$$

Как отмечалось ранее множество адресов является подмножеством декартова произведения множеств значений реквизитов $A \subset R_1 \times R_2 \times \dots \times R_n$. Поэтому естественно ожидать, что множество нормальных адресов принадлежит декартовому произведению множеств нормальных значений реквизитов.

R_{bi} назовем базовым подмножеством всех значений реквизита R_i $i \in [1, n]$ для которых справедливо:

$$\forall r_{bi} \in R_{bi} \exists a_b \in A_b \wedge a_b \in R_1 \times R_2 \times \dots \times R_{i-1} \times r_{bi} \times R_{i+1} \times \dots \times R_n \quad (3)$$

Т.е. множество базовых значений каждого реквизита составляют значения реквизитов, которые присутствуют в составе множества базовых адресов.

R_{ci} назовем подмножеством аннулированных значений реквизита R_i $i \in [1, n]$ если

$$\forall r_{ci} \in R_{ci} \rightarrow \forall r_{ci} \notin R_{bi} \wedge$$

$$\exists a_c \in A_c \wedge a_c \wedge R_1 \times R_2 \times \dots \dots \times R_{i-1} \times r_{ci} \times R_{i+1} \times \dots \times R_n \quad (4)$$

Т.е. множество аннулированных значений каждого реквизита составляют значения реквизитов, которые присутствуют в составе множества аннулированных адресов, но отсутствуют в множестве базовых адресов реквизита.

Таким образом, множеством **нормальных значений реквизита** называется объединение подмножеств базовых и аннулированных значений реквизита.

$$\bar{R} \equiv R_b \cup R_c \quad (5)$$

Из определения множества нормальных значений реквизита следует, что пересечение множеств базовых и аннулированных значений каждого реквизита пусто.

$$R_b \cap R_c = \emptyset \quad (6)$$

Это свойство следует из определения множества аннулированных значений реквизита.

Множество базовых адресов принадлежит декартовому произведению множеств базовых значений реквизитов

$$A_b \subseteq R_{b1} \times R_{b2} \times \dots \times R_{bi} \times \dots \times R_{bn} \quad (7)$$

Действительно, если существует базовый адрес, который не принадлежит декартовому произведению множеств базовых значений реквизитов, т.е.

$$\exists a_b \in A_b \wedge a_b = \{r_1, r_1, \dots, r_n\} \rightarrow r_{bi} \notin R_{bi} \quad (8)$$

Но такое предположение вступает в противоречие с тем требованием, что R_{bi} содержит **все значения** этого реквизита, которые содержатся в базовых адресах.

Множество аннулированных адресов принадлежит следующему объединению декартовых произведений множеств реквизитов

$$A_c \subseteq \left\{ \begin{array}{l} R_{c1} \times \bar{R}_2 \times \dots \times \bar{R}_n; \\ \bar{R}_1 \times R_{c2} \times \dots \times \bar{R}_n; \\ \dots \\ \bar{R}_1 \times \bar{R}_2 \times \dots \times R_{ci} \times \bar{R}_n; \\ \dots \\ \bar{R}_1 \times \bar{R}_2 \times \dots \times R_{cn}. \end{array} \right. \quad (9)$$

Т.е. множество аннулированных адресов содержится в объединении множеств, составленных из декартовых произведений значений реквизитов. Где каждое произведение содержит множество аннулированных значений хотя бы для одного реквизита и нормальные множества для остальных реквизитов.

Совокупность свойств (12) и (14) образует необходимый **признак нормальности множества адресов \bar{A}** .

2, Нормализация значений реквизитов

Реальное множество значений каждого реквизита шире чем множество его нормальных значений. Поэтому рассмотрим несколько шагов расширения множества нормальных значений реквизита в направлении множества реальных значений реквизита. На каждом шаге будет рассматриваться правило (функция) нормализации расширенного значения реквизита, т.е. преобразование его к одному из нормальных значений.

Шаг 0 - синтаксическая нормализация

Значения реквизита могут иметь форму имени или форму номера.

Требования к синтаксису значений реквизита в форме имени.

1. Имя не должно начинаться и оканчиваться символом «пробел»;
2. Отдельные слова имени должны быть разделены только одним символом «пробел»;
3. Между символом «дефис» и составными частями слов в имени не должно быть символа «пробел»;
4. Между открывающей скобкой и последующим за ней словом в имени не должно быть символа «пробел»;
5. Между закрывающей скобкой и предшествующим ей словом в имени не должно быть символа «пробел»;
6. Между словом следующей за ним точкой не должно быть символа «пробел».

Требования к синтаксису значений реквизита в форме номера.

1. Номер не должно начинаться и оканчиваться символом «пробел»;
2. Номер начинаться с цифры;
3. Число в номере может оканчиваться одной буквой, но между числом и буквой не должно быть ни одного символа «пробел»;
4. Буква, завершающая число, не должна быть заключена ни в одинарные ни в двойные кавычки;
5. Номер может содержать не более 2-х чисел, которые отделены друг от друга символом «/»;
6. Между символом «/» и окружающими его числами не должно быть символов «пробел».

0-расширение значений реквизита

0-расширение значений реквизита R_{ex0} – это множество значения реквизита, которое приводятся к нормальному значению путем исправления одного или нескольких недостатков имени, перечисленных в требованиях к значению в форме имени или номера.

$$R_{ex0} \cap \bar{R} = \emptyset \quad (10)$$

Множеством **0-расширенных значений реквизита** будем называть объединение нормальных значений реквизита и 0-расширения значений этого реквизита

$$\bar{R}_{ex0} \equiv R_{ex0} \cup \bar{R} \quad (11)$$

0-нормализацией или функцией 0-нормализации ν_0 называется функция из R_{ex0} в \bar{R} или $R_{ex0} \xrightarrow{\nu_0} \bar{R}$ такая, что:

$$\nu_0(r) \equiv \begin{cases} r, \forall r \in \bar{R} \\ \forall r \in R_{ex0} \exists r' (r' \in \bar{R}) \nu_0(r) = r' \end{cases} \quad (12)$$

Другими словами, функция ν_0 оставляет неизменными нормальные значения реквизита, а каждому значению 0-расширения ставит в соответствие нормальное значение этого реквизита.

Шаг 1 - орфографическая нормализация

Требования к орфографии значения реквизита в форме имени.

1. Написание всех слов имени или номера должно в точности соответствовать их написанию одному из 0-расширенных значений реквизита.

1-расширение значений реквизита

1-расширение значений реквизита R_{ex1} – это множество значений реквизита, каждое слово в котором отличается от 0-расширенного значения реквизита значения реквизита не больше чем на 1 по метрике Дамерау-Левенштейна. При этом значения 1-расширения не принадлежат $\overline{R_{ex0}}$.

$$R_{ex1} \cap \overline{R_{ex0}} = \emptyset \text{ и} \quad (13)$$

Множеством **1-расширенных значений реквизита** будем называть объединение 0-расширенных значений и 1-расширения значений реквизита

$$\overline{R_{ex1}} \equiv R_{ex1} \cup \overline{R_{ex0}} \text{ и} \quad (14)$$

1-нормализацией или функцией 1-нормализации v_1 называется функция из $\overline{R_{ex1}}$ в $\overline{R_{ex0}}$ или $\overline{R_{ex1}} \xrightarrow{v_1} \overline{R_{ex0}}$ такая, что:

$$v_1(r) \equiv \begin{cases} r, \forall r \in \overline{R_{ex0}} \\ \forall r \in R_{ex1} \exists r' (r' \in \overline{R_{ex0}}) v_1(r) = r' \end{cases} \quad (15)$$

Другими словами, функция v_1 оставляет неизменными 0-расширенные значения реквизита, и каждому значению 1-расширения ставит в соответствие 0-расширенное значение реквизита.

Шаг 2 – нормализация следования

Требования к следованию слов в значении реквизита.

1. Порядок слов в значении реквизита должно в точности соответствовать их порядку одного из 1-расширенных значений реквизита.

2-расширение значений реквизита

2-расширение значений реквизита R_{ex2} – это множество значений реквизита, которые получены путем перестановок слов в значениях множества 1-расширенных значений реквизита, но не принадлежащие R_{ex1}

$$R_{ex2} \cap \overline{R_{ex1}} = \emptyset \quad (16)$$

Множеством **2-расширенных значений реквизита** будем называть объединение множества 1-расширенных значений реквизита и 2-расширения значений

$$\overline{R_{ex2}} \equiv R_{ex2} \cup \overline{R_{ex1}} \quad (17)$$

2-нормализацией или функцией 2-нормализации v_2 называется функция из $\overline{R_{ex2}}$ в $\overline{R_{ex1}}$ или $\overline{R_{ex2}} \xrightarrow{v_2} \overline{R_{ex1}}$ такая, что:

$$v_2(r) \equiv \begin{cases} r, \forall r \in \overline{R_{ex1}} \\ \forall r \in R_{ex2} \exists r' (r' \in \overline{R_{ex1}}) v_2(r) = r' \end{cases} \quad (18)$$

Другими словами, функция v_2 оставляет неизменными 1-расширенные значения реквизита, и каждому значению 2-расширения ставит в соответствие 1-расширенное значение, т.е. приводит порядок слов значения реквизита 2-рас-

ширения в соответствие с порядком слов соответствующего значения множества 1-расширенных значений этого реквизита.

Шаг 3 – смысловая нормализация

Требования однозначности смыслового значения реквизита.

Значение реквизита r должно быть однозначным, т.е. должно отсутствовать другие значения реквизита, определяющих ту же самую территорию, а, следовательно, множество объектов.

$$FT(r) = t \wedge \forall r' (r' \in R) FT(r') = t \rightarrow r' = r \quad (19)$$

3-расширение адресов

3-расширение значений реквизита R_{ex3} – это множество значений, не входящих в R_{ex2} , но определяющих те же территории, что и значения R_{ex2} .

$$R_{ex3} \cap \overline{R_{ex2}} = \emptyset \quad (20)$$

Более точно

$$\forall r' (r' \in R_{ex3} \wedge r' \notin \overline{R_{ex2}}) \exists r (r \in \overline{R_{ex2}}) FT(r') = FT(r) \quad (21)$$

Множеством **3-расширенных значений реквизита** будем называть объединение множества 2-расширенных значений реквизита и 3-расширения значений

$$\overline{R_{ex3}} \equiv R_{ex3} \cup \overline{R_{ex2}} \quad (22)$$

Пусть $r (r \in \overline{R_{ex2}})$, тогда $r' (r' \in R_{ex3}) FT(r') = FT(r)$ – называется **синонимом** r . А r в свою очередь будем называть исходным значением реквизита.

3-нормализацией или функцией 3-нормализации v_3 называется функция из $\overline{R_{ex3}}$ в $\overline{R_{ex2}}$ или $\overline{R_{ex3}} \xrightarrow{v_3} \overline{R_{ex2}}$ такая, что:

$$v_3(r) \equiv \begin{cases} r, \forall r \in \overline{R_{ex2}} \\ \exists r' (r' \in \overline{R_{ex3}}) v_3(r) = r' \end{cases} \quad (23)$$

Другими словами, функция v_3 оставляет неизменными 2-расширенные значения реквизита, и каждому значению 3-расширения ставит в соответствие 2-расширенное значение, т.е. каждому синониму ставит в соответствие исходное значение реквизита.

Замечание 1.

Последовательное применение к синониму r' всех функций нормализации приводит к получению нормального значения реквизита, принадлежащего множеству нормальных значений реквизита \overline{R} . Другими словами,

$$\forall r' (r' \in \overline{R_{ex3}}) v_0 (v_1 (v_2 (v_3 (r')))) \in \overline{R} \quad (24)$$

Конец замечания.

3. Нормализация адресов

Реальное множество адресов шире чем множество нормальных адресов. В первую очередь из-за возможности расширения значений реквизитов. Кроме того, расширение множества адресов за пределы нормального множества может возникать независимо от того нормализованы составляющие его значения реквизитов или нет.

Первые 4 шага расширения множества нормальных адресов в направлении множества реальных адресов – это шаги расширения множеств значений реквизитов, из которых составлены адреса. Как и в случае с реквизитами, на каждом шаге будет рассматриваться правило (функция) нормализации расширенного адреса, т.е. преобразование его к одному из нормальных адресов.

Пятый дополнительный шаг расширения множества нормальных адресов характерен только для полного адреса, как совокупности всех значений, составляющих его реквизитов.

Шаг i – нормализация на основе реквизитов

Шаги нормализации адресов с номерами от 0 до 3 – это шаги на основе нормализации значений реквизитов, составляющих адрес. Для краткости вместо описания каждого шага дадим общее описание этих шагов, заменив номер шага символом i .

i -расширение адресов A_i – это множество адресов, значения хотя бы одного из составляющих его реквизитов подвергнуты i -расширению.

Множеством i – **расширенных адресов** будем называть объединение множества i **1-расширенных адресов** и i -расширения адресов

$$\overline{A_i} \equiv A_i \cup \overline{A_{i-1}} \quad (25)$$

i -нормализацией или функцией i -нормализации v_i называется функция из A_i в $\overline{A_{i-1}}$ или $\overline{A_i} \rightarrow \overline{A_{i-1}}$, выполняющая i -нормализацию значений каждого реквизита, составляющего адрес что:

$$v_i(a) \equiv \begin{cases} a, \forall a \in \overline{A_i} \\ \exists a'(a' \in \overline{A_{i-1}})v_i(a) = a' \end{cases} \quad (26)$$

Другими словами, функция v_i оставляет неизменными адреса из множества i – **1-расширенных** адресов, и каждому адресу i -расширения ставит в соответствие i – **1-расширенный** адрес.

Шаг 4 - смысловая нормализация адресов

Требования однозначности смыслового значения адреса.

Адрес a должен быть **однозначным**, т.е. должны отсутствовать адреса, определяющие тот же самый объект недвижимости объектов.

4-расширение адресов

4-расширение адресов A_4 – это множество адресов, не входящих в $\overline{A_3}$, но определяющих те же те же объекты недвижимости, что и значения $\overline{A_3}$.

$$A_4 \cap \overline{A_3} = \emptyset \quad (27)$$

Более точно

$$\forall a'(a' \in A_4 \wedge a' \notin \overline{A_3}) \exists a (a \in \overline{A_3})F(a') = F(a) \quad (28)$$

Множеством **4-расширенных значений адресов** будем называть объединение множества 3-расширенных значений адресов и 4-расширения значений

$$\overline{A_4} \equiv A_4 \cup \overline{A_3} \quad (29)$$

Пусть $a(a \in \overline{A_3})$, тогда $a'(a' \in R_4)F(a') = F(a)$ – называется **синонимом a** . А a в свою очередь будем называть исходным адресом.

4-нормализацией или функцией 4-нормализации v_4 называется функция из A_4 в $\overline{A_3}$ или $\overline{A_4} \rightarrow \overline{A_3}$ такая, что:

$$v_4(a) \equiv \begin{cases} a, \forall a \in \overline{A_3} \\ \exists a'(a' \in \overline{A_3})v_4(a) = a' \end{cases} \quad (30)$$

Другими словами, функция v_4 оставляет неизменными 3-расширенные адреса, и каждому значению 4-расширения ставит в соответствие 3-расширенный адрес, т.е. каждому синониму адреса ставит в соответствие исходный адрес.

Замечание 1.

Последовательное применение к синониму a' всех функций нормализации приводит к получению нормального адреса, принадлежащего множеству нормальных адресов \overline{A} . Другими словами,

Конец замечания.

Конфликт интересов отсутствует.
There is no conflict of interest.

СПИСОК ЛИТЕРАТУРЫ

1. **Фреге Готлоб**, О Смысле и Значении, стр. 230-246. Логика и логическая семантика: Сборник трудов/Пер. с нем. Б.В. Бирюкова под ред. З.А. Кузичевой: Учебное пособие для студентов вузов. М.: Аспект Пресс. 2000 – 512с.
2. **Бах Эммон**, Неформальные лекции по формальной семантике: Пер. с англ. / Под ред. О. А. Митрофановой, О. В. Митрениной; Предисл. Б. Парти, - М.: Книжный дом «ЛИБРОКОМ», 2010.- 224 с.
3. **Стаффорд Бир**. Мозг фирмы. Пер. с англ. – М.: Издательство: Едиториал УРСС, 2005. 416с.
4. **Нагаев Р.Т.** Недвижимость: Энциклопедический словарь. – Казань: Издательство ГУП «Идел- Пресс», 2003. 1088-1.
5. Адрес объекта недвижимости. Словари и энциклопедии на Академике. https://official.academic.ru/479/Адрес_объекта_недвижимости
6. Единые требования к описанию адресов при ведении ведомственных информационных ресурсов(утв. приказом Федеральной налоговой службы от 31 августа 2011 г. № ММВ-7-1/525@) ГАРАНТ.РУ: <http://www.garant.ru/products/ipo/prime/doc/12089454/#ixzz53OX7qDTV>
7. ГК РФ Статья 130. Недвижимые и движимые вещи. http://www.consultant.ru/document/cons_doc_LAW_5142/2a54eb7c0c87a49c41aa10efb253f6bdea2bfcf4/
8. Правила присвоения, изменения и аннулирования адресов (утв. постановлением Правительства РФ от 19 ноября 2014 г. N 1221) Система ГАРАНТ: <http://base.garant.ru/70803770/>
9. Старые и современные названия улиц Красноярска, <http://www.krasplace.ru/starye-i-sovremennye-nazvaniya-ulic-krasnoyarska>
10. **Шефель О. М:** Первая полная энциклопедия ЖКХ. – Москва: Издательство АСТ, 2013. 544. <http://ast.ru/catalog/23040/828296.php>

11. Положение о порядке присвоения, изменения и аннулирования адресов объектам недвижимости в городе Красноярске, утвержденное Постановлением администрации города от 20 апреля 2015 года N 236. <http://www.admkrsk.ru/sites/doc/Pages/detail.aspx?RecordID=14919>
12. **Гладков С.Л.** О требованиях к интеллектуальной модели данных. ВЕСТНИК Московского Университета имени С.Ю.Витте, 2015>2(10). – С. 63-70, http://www.muiv.ru/vestnik/pdf/pp/ot_2015_2_63-70.pdf
13. **Гладков С.Л.** Формализация понятия адрес

Поступила 06.09.2018
УДК 004.891:510.63

ИНФОРМАЦИЯ ОБ АВТОРЕ

Гладков Сергей Львович – технический директор ООО «Айгео»: e-mail: gladkovs@list.ru

AUTHOR'S INFORMATION

Sergey L'vovich Gladkov – technical director IGEO Limited
e-mail: gladkovs@list.ru