IOP Conference Series: Materials Science and Engineering

PAPER • OPEN ACCESS

Using the set-theoretic approach to formalize the concept of address

IOP Conf. Series: Materials Science and Engineering **537** (2019) 052004 doi:10.1088/1757-899X/537/5/052004

Using the set-theoretic approach to formalize the concept of address

G M Rudakova[1], S L Gladkov[2] and O V Korchevskaya[1]

[1] Reshetnev Siberian State University of Science and Technology, 31, Krasnoyarsky Rabochy Av., Krasnoyarsk, 660037, Russian Federation
[2] LLC Aigeo, 105, office 4-34, 60 years of October ul., Krasnoyarsk, 660079, Russian Federation

E-mail: gmfait@gmail.com

Abstract. The paper deals with the problem of using the address as an identifier for the location of the property. A formal description of the common notion of “Address” (building address, citizen registration address, etc.) has been introduced, which can be used in information systems and geo-referencing algorithms. In the article, the address is analyzed as a set of values of details. A formal definition of address is formulated in the space of attribute values. The formal semantics of the introduced concept is presented. The principles are formulated, allowing to use the address as an identifier of objects in databases. The source materials for the article are federal and regional regulations.

1. Introduction

The problem of unambiguous identification of objects of any nature is quite relevant for various spheres of society. As an example, the problems of identifying objects and subjects on the Internet, negative factors of the working environment, automatic identification of barcode commodity codes, addressing of real estate objects, citizens' registration, or, for example, the email address of a product.

Subject areas are quite diverse, have their own specifics. The concept of an address as the identifier of the property location is taken as the main object of consideration.

The need to introduce a strict formalization of the concept of address is associated with the problems of its use in information systems, geolocation algorithms. The paper uses the set-theoretic and semantic approaches. The paper uses the set-theoretic and semantic approaches [1 – 3].

2. Address as a concept

From the analysis of the definitions of the address [4], it follows that the address is a point in the space of requisites:

$$a=\{r_1,r_2,\ldots r_n\}; \tag{1}$$

where n is the number of requisites, r_i is the value of i-th requisite $1\leq i\leq n$.

Then the set of admissible addresses A is a subset of the Cartesian product of sets of requisites:

$$A\subset R_1\times R_2\times\ldots\times R_n\,; \tag{2}$$

where n is the number of requisites, R_i – set of values of the i-th requisite, $1\leq i\leq n;$

or

$$a \in A \leftrightarrow r_1 \in R_1 \wedge r_2 \in R_2 \wedge \ldots \wedge r_n \in R_n. \quad (3)$$

In addition, each address determines the location of the addressing object [5]. This address property can be described as follows.

Let O be the set of addressing objects, o an element of the set O, that is, $o \in O$.

Let, also, F be a function that displays a set of addresses on a set of addressing objects:

$$A \xrightarrow{F} O. \quad (4)$$

Then the binding address for each addressing object, that is, the existence of an address for any addressing object, is determined as follows:

$$F(A)=O \text{ or } \forall o\ (o \in O) \rightarrow \exists a\ (F(a)=o). \quad (5)$$

The uniqueness of the address, i.e. the uniqueness of the addressing object corresponding to one address is defined as

$$\forall a\ \exists o_1,o_2 (a \in A \wedge F(a)=o_1 \wedge F(a)=o_2) \rightarrow (o_1=o_2)\ (a). \quad (6)$$

From properties (5) and (6) it follows that at least one address corresponds to each addressing object. If there is an addressing object corresponding to the address, then it is unique.

3. Address semantics

The article [6] proposes to consider the concepts in a pair with their context, where the context is a logical statement linking the generating concepts with their meanings.

In the case of addresses, the generating concepts are the address details, therefore, the address coincides with its own context C (a), i. e. the following is true:

$$C(a) \equiv a \equiv \left(R_1 = r_1) \wedge (R_2 = r_2) \wedge \ldots \wedge (R_n = r_n\right). \quad (7)$$

The sequence $r_1, r_2, \ldots, r_n$ is called the address protocol.

Mandatory and unique properties allow addressing to be considered as the context of the addressing object:

$$C(o)=a. \quad (8)$$

We emphasize that the context of the object of addressing is not necessarily uniquely the object of addressing. Context is a general characteristic of the set of addressing objects, as well as other concepts. Therefore, we can consider the context of each address requisite. In other words, each requisite of the address R_i has its own context.

$$C(R_i) \equiv \left(R_1 = r_1) \wedge (R_2 = r_2) \wedge \ldots \wedge (R_{i-1} = r_{i-1}\right); \quad (9)$$

where $2 \le i \le n$.

$$C(R_1) \equiv true. \quad (10)$$

4. Requisite as a concept

The set of values of each attribute, as well as the set of addresses, has an internal structure, since the whole set of values of the attribute R is divided into m disjoint subsets:

$$R = R_{t1} \cup R_{t2} \cup \ldots \cup R_{tm},\ \forall i,j (i,j \in [1,m]) R_{ti} \cap R_{tj} = \emptyset. \quad (11)$$

Each subset R_{ti} corresponds to a certain type of requisite values. Therefore, it is more convenient to represent the values of r of the attribute of R as a pair by the type of the value of the attribute and the immediate value within the subset corresponding to this type:

$$r=\{tr,vr\},\ r\in R. \tag{12}$$

The set of allowable values for props R is a subset of the Cartesian product of sets of value types and props values themselves:

$$R\subset TR\times VR; \tag{13}$$

or

$$r\in R\leftrightarrow tr\in TR\wedge vr\in VR. \tag{14}$$

In addition, each requisite determines the location of the territories in which the properties are located. This property of the address can be described as follows:

Let O be a set of addressing objects, o – element of set O, i. e. $o\in O$, T – set of territories, t – element of set T, i e. $t\in T$.

Let, also, FT be a function to each value of the requisite sets in correspondence one territory from a set of territories, and TO - a display, which assigns to each territory a set of objects of addressing:

$$R\overset{FT}{\rightarrow}T\overset{TO}{\rightarrow}O. \tag{15}$$

From the property of the mandatory address for each addressing object, the existence of the value of the attribute and the territory for any addressing object follows, is determined as follows:

$$\forall o\ (o\in O)\rightarrow\ \exists t\in T,r\in R\ \left(FT(r)=t\wedge o\in TO(t)\right). \tag{16}$$

From properties (15) and (16) it follows that each addressing object corresponds to at least one territory, on which at least this addressing object and one value of the attribute, which is part of the address, is located.

5. Set of Addressing objects

Let O be the set of addressing objects, and R_1-the first requisite, i.e. requisite without context ($C(R_1)\equiv true$). Then from the property of the mandatory address follows, that set $\{TO{*}FT(r_1)\}$, where r_1 arbitrary value of requisite forms a covering of the set O. If this was not the case, then the address binding property would be violated, since for objects from $O/\{TO{*}FT(r_1)\}$, there is no address, because there is no value for the requisite R_1, which defines this subset:

$$\cup_{r_1} TO{*}FT(r_1)=O. \tag{17}$$

Let R_2 be the second requisite, i.e. $C(R_2)\equiv(R_1=r_1)$, then:

$$\cup_{r_2} TO{*}FT(r_2)=TO{*}FT(r_1)\subset O. \tag{18}$$

In other words,

$$\forall r_2\in R_2\ TO{*}FT(r_2)\subset TO{*}FT(r_1). \tag{19}$$

The proof is obvious. The set $TO{*}FT(r_1)$ is on the one hand a subset of O, on the other hand, the same subset is determined by the context $C(R_2)\equiv(R_1=r_1)$, i.e. determines the set of all values of the superposition of functions $TO{*}FT(r_2)$.

The set $\{TO{*}FT(r_2)\}$ is covering the set $TO{*}FT(r_1)$.

6. Legitimacy, relevance, uniqueness of addresses

Until now, the relevance and legitimacy of address a was implicitly viewed as synonymous with its belonging to the set of addresses A:

$$\mathrm{s}(a)=\begin{cases}legitimate,\ a\in A\\ missing,\ a\notin A;\end{cases} \tag{20}$$

where $s(a)$ is address status determination function.

Such a model is non-constructive in conditions when it is necessary to take into account obsolete, renamed addresses, as well as addresses with deformed values of requisites, i.e. addresses containing syntax errors.

Consider a set of addresses consisting of a combination of legitimate and canceled addresses:

$$\overline{A} \equiv A_l \cup A_c; \tag{21}$$

where $A_l \equiv A$ is a set of legitimate addresses, A_c- a set of revoked (canceled) addresses. Sets of legitimate and canceled addresses do not overlap $A_l \cap A_c = \emptyset$.

The function of the address status on set $\overline{A} \equiv A_l \cup A_c$ is:

$$\mathrm{s}(a) = \begin{cases} legitimate,\ a \in A_l \\ cancelled,\ \ a \in A_c. \\ missing,\ a \notin \overline{A} \end{cases} \tag{22}$$

The relevance of the address implies the possibility of its verification. This feature is provided by the previously introduced function *F*, which displays a set of addresses to a set of addressing objects:

Then the relevance of the address is defined as the presence of an addressing object, whose location determines this address:

$$a\text{-}relevant \Leftrightarrow \exists o\ (o \in O)\ F(a) = o. \tag{23}$$

And, finally, the actual address of *a* is unambiguous, if there are no other addresses defining the location of the object *o*:

$$F(a) = o \wedge \forall a_i (a_i \in A_l) F(a_i) = o \longrightarrow a_i = a. \tag{24}$$

7. Sign of uniqueness of the address

With a suitable choice of requisites, the address uniquely identifies the location of the property if and only if each attribute contains unique values in the range of acceptable values.

Each value of the previous address requisite corresponds to a set of values or a dictionary of the next address attribute. Therefore, each address attribute, except the oldest one, is a ratio of the set of values of the previous requisite and dictionaries of the values of the current requisite:

$$R_i \subset R_{i-1} \times DR_i; \tag{25}$$

where DR_i is the combined dictionary of meanings $(2 \leq i \leq n)$.

The combined dictionary of values is the union of dictionaries corresponding to the values of the previous requisite.

$$DR_i \equiv \bigcup_{j=1}^{n} DR_{i,j}\ ; \tag{26}$$

where $DR_{i,j}$ is a dictionary of values corresponding to the *j*-th element of the dictionary of the previous requisite.

Further, DR_1 and R_1 will be considered synonymous in order not to write complex formulas such as:

$$\begin{cases} r_j \in DR_{i-1},\ i > 2 \\ r_j \in R_1, i = 2. \end{cases} \tag{27}$$

Each requisite of the address uniquely defines its own subset of address-forming elements. That is, for each requisite there is a function that maps its values to a subset of the addressing objects:

$$\forall i (1 \leq i \leq n)\ R_i \xrightarrow{FR_i} OR_i\ ; \tag{28}$$

where $OR_i \subset O$.

IOP Conf. Series: Materials Science and Engineering **537** (2019) 052004 doi:10.1088/1757-899X/537/5/052004

In this case, the uniqueness of the values of the requisites is that the values of each requisite are converted by the function into a set of disjoint subsets of the addressing objects:

$$\forall i\,(1 \le i \le n)\,\forall(1 \ge j \ge m)\,\bigcap_{j=1}^{m} FR_i\left(r_{i,j}\right) \equiv \emptyset; \tag{29}$$

where m is the number of elements in dictionary R_i.

The dependence of the values of the details is that the set of objects of addressing, corresponding to the value of the junior requisite, belongs to the set of objects of addressing, corresponding to the value of the higher attribute:

$$\forall i\,(2 \le i \le n)\,\forall(1 \ge j \ge m)\,\forall(1 \ge k \ge l) \quad FR_{i-1}\left(r_{i-1,j}\right) \supset FR_i\left(\left(r_{i-1,j}, r_{i,k}\right)\right); \tag{30}$$

where m – number of elements in the dictionary R_{i-1}, a l – number of elements in the dictionary R_i.

8. Conclusion

The article discusses the possibility of using the address as an identifier of the property. At the same time, obligation and uniqueness allow to consider the address as the context of the addressing object.

A formal representation of the concept of address as a set of values of attributes is formulated. The set-theoretic approach, the algebra of logic, are used. Also addressed issues of legitimacy, relevance, unique addresses.

Further research is aimed at bringing an arbitrary address to normal form, developing algorithms for the normalization of details and addresses, as well as the use of this approach in other subject areas.

References

[1] Gottlob F 2000 *Logic and Logical Semantics* (Moscow: Aspect Press)
[2] Emmon B 2010 *Informal Lectures on Formal Semantics* (Moscow: Book House LIBROKOM)
[3] Beer S 2005 *Brain Of The Firm* (Moscow: Publisher: Editorial URSS)
[4] Nagayev RT 2003 *Real Estate: Encyclopedic Dictionary* (Kazan: Publisher GUP «Idel-Press»)
[5] Scheffel O M 2013 *First Full Encyclopedia of Housing and Communal Services* (Moscow: Publisher AST)
[6] Gladkov S L 2015 On the requirements for an intelligent data model *BULLETIN of Moscow University named after Witte S Yu* **2 (10)**