Formalizing Conceptual Spaces

Martin RAUBAL
Institute for Geoinformatics, University of Münster
Robert-Koch-Str. 26-28, 48149 Münster, Germany
raubal@uni-muenster.de

Abstract. Semantics constitutes the highest level on the communication ladder. The mutual understanding of terms makes it possible that information systems can communicate with people and among each other. Two opposing approaches to explaining the meaning of terms exist: a realist and a cognitive one. Many efforts for solving the so-called problem of semantic interoperability in the area of Information Science are based on realist semantics, which claims that meaning is in the world. In this paper we argue for a cognitive approach to semantic interoperability, which is based on the assumption that meanings are in the heads of people. This allows us to account for the fact that different people have different conceptualizations of the world and therefore require different answers and presentations of answers to their spatio-temporal questions. The paper presents a formal approach for representing Gärdenfors' idea of conceptual spaces—sets of quality dimensions within a geometrical structure. It complements the mathematical notion of a vector space with a standardization method from statistics to formally define *conceptual vector spaces*. Such spaces allow for the measurement of semantic distances between instances of concepts and also for the assignment of weights to their quality dimensions in order to represent different contexts. Mappings, such as transformations and projections, between such spaces facilitate knowledge sharing and therefore support *cognitive semantic interoperability*. A case study from the geospatial domain—wayfinding services with landmarks—demonstrates the usefulness and plausibility of the approach.

Introduction

Communication or the activity of conveying information depends on semantics, i.e., the meanings of words in a language[1]. In previous times the exchange of information was between human beings, whereas now information gets additionally exchanged between various computer (information) systems, and between such systems and human beings. In order to provide for a successful exchange of information in the sense that the content is understood[2] in an unambiguous and consistent way by different parties, their semantics needs to be defined and represented in a formal way.

Two opposing approaches to explaining the meaning of terms exist: a *realist* and a *cognitive* one. Realist semantics asserts that the meanings of expressions are in the world and therefore independent of how individual people understand them. In cognitive semantics, meanings are mental entities, i.e., mappings from expressions to conceptual structures, which themselves refer to the real world. It has been demonstrated that realist approaches to semantics face a number of difficulties, most notably their problems to deal with learning, with mentally constructed objects that have no direct correspondence in the real world,

[1] We refer to language in a general sense including both natural and computer languages.
[2] This means human *understanding* and its metaphorical projection to computers.

and with the fact that the meaning of concepts often changes both over time and between contexts [1].

When creating automated information services—services that work with each other without human intervention—it is often forgotten that the final goal is to create information for a human user. Therefore it is important how individual people understand the output of a computer system and that the computer system understands people's meanings of terms used to formulate a question to the system. In this paper we present a formal method to represent *conceptual spaces*—sets of quality dimensions within a geometrical structure—as proposed by Gärdenfors [1]. Computer-friendly representations in the area of cognitive semantics are rare but urgently needed. They allow us to account for the fact that different people have different conceptualizations of the world and therefore require different presentations of answers to their spatio-temporal queries. The concepts of system designers have to be matched with the concepts of individual users to make communication a successful enterprise.

The way of formalizing conceptual spaces presented here utilizes the mathematical theory of vector spaces in combination with a method from statistics. Concepts that are represented through conceptual spaces are usually built by combining different domains. For example, one's concept of a building may combine the domains of size, shape, and color. A concept is represented as an n-dimensional conceptual vector space, whose axes represent quality dimensions or further vector spaces denoting domains. In order to standardize the variables represented by these axes a statistical method called z-transformation is applied [2, 3]. This allows for representing instances of a concept (i.e., individuals) as points in the conceptual vector space and measure the distances between them and their distances to a prototype of the concept. Such prototype is ideally an n-dimensional region in the conceptual space [1]. In addition, one can account for the use of a concept in different contexts by assigning different weights to the axes of the conceptual vector space.

Formal conceptual spaces can be exploited for the matching of concepts between parties of a communication process, e.g., for the communication between a computer system and a user. A particular concept used by the system is represented through a conceptual space (i.e., the designer's) and the same concept is represented through the user's conceptual space. Both spaces can then be compared—for example, by representing one's prototype within the other space and determining the distance between the prototypes—and adjusted to each other. For practical reasons, the system adapts the semantics of its concepts to the user's semantics, leading to improved human-computer interaction. In order to demonstrate how the formalization of conceptual spaces works in practice we use a case where a navigation service gives instructions to a wayfinder. These instructions include facades of buildings as landmarks [4]. The concept of 'facade' is represented in two conceptual spaces, both from the system's and the user's perspective.

1 Cognitive Semantics and Conceptual Spaces

Cognitive semantics claims that the meanings of terms are in people's heads. As such, meanings are mappings to conceptual structures, which themselves refer to real-world entities. Cognitive semantics tries to give answers to many of the problems a realist semantics account of reality faces, such as explaining processes of learning and the construction of mental objects that do not correspond to real-world features. The realist approach to semantics assumes that meaning consists only of relationships between abstract symbols and elements in real-world models[3]. Therefore, correct reasoning is achieved by logical manipulation of such symbols and elements. This point of view lacks a place for people, because according to realist semantics the world stays the same, whether there are people in it or not [5]. But information systems, which can interact with their users in a comprehensible way, require different mental knowledge representations of different users [6]. As Rosch puts it: "It should be emphasized that we are talking about a perceived world and not a metaphysical world without a knower" [7, p.29].

Lakoff [5] argues that people use cognitive models in trying to understand the world. He stands behind the experientialist view, which claims that reason is made possible by the body. Language makes use of our cognitive system and we organize our knowledge by means of structures. Structure plays an important role in cognitive semantics [8]. The basic conceptual units of cognitive semantics are internally structured and there exist meaningful relationships between their components. Prime examples for such a conceptual unit are image schemata—recurring mental patterns that help people to comprehend and structure their daily experiences [9]. An image schema can be seen as a very generic, maybe even universal, and abstract structure that helps people establish a connection between different experiences that have this same recurring structure in common. The internal structures of these units make it possible to establish mappings across conceptual models and therefore between domains, e.g., metaphors [10] and blended spaces [11][4]. Such mappings belong to the distinguished capabilities of human cognition with regard to establishing and communicating meaning [13].

Gärdenfors [1] argues that cognitive science needs three levels, i.e., the symbolic, the conceptual, and the subconceptual level. His major endeavor is the explanation of the conceptual level, which consists of conceptual spaces. According to Gärdenfors, a conceptual space is a set of quality dimensions with a geometrical or topological structure for one or more domains. A domain is represented through a set of integral dimensions, which are distinguishable from all other dimensions. For example, the color domain is formed through the dimensions hue, saturation, and brightness. On the conceptual level, learning corresponds to the extension of a conceptual space by new quality dimensions. Every object is represented as a point in a conceptual space. The similarity of two objects can therefore be expressed as the distance between their points in the conceptual space.

Our meanings of concepts can change over time and they vary depending on the context in which they are used. Conceptual spaces are capable of representing such dynamic aspects by giving different saliencies to dimensions and domains [1]. So far, only a few attempts to formalize elements of cognitive semantics have been made. For example, some formal representations of image schemata exist [14-17]. There is a need for formalized conceptual spaces so that they can be implemented in order to facilitate knowledge sharing [1] and communication between computer systems, and between systems and their users. A general formulation of conceptual spaces can be found in [18]. There, conceptual spaces are represented as pointed metric spaces, which generalize the commonly used vector spaces. In their conclusions, the authors argue that there has been cognitive support for the position that dimensions are quantitative and addition should be defined. This paper presents a proposal in this direction by using vector spaces, which allow for addition and scalar multiplication.

[3] For the distinction between extensional realist semantics and intensional realist semantics, where terms are mapped to sets of possible worlds, see Gärdenfors [1].

[4] A practical example for deriving the semantics of geographic categories through conceptual integration is given by Kuhn [12].

2 Formalization Methods

This section describes the formal methods employed to formalize conceptual spaces and calculate distances between instances of concepts in them.

2.1 Vector Space

Vector spaces are of central importance for the mathematical subfield of linear algebra [19, 20]. A vector space $\mathbf{R}^n$ consists of all column vectors with n components, where the components are real numbers. For example, in $\mathbf{R}^3$ the three components give a point in three-dimensional space. Within all vector spaces one can add any two vectors and multiply vectors by scalars. Formally, a real vector space is a set of vectors with rules for vector addition and multiplication by real numbers, such as associativity, commutativity, etc.[5]. This formal definition also allows for vectors being matrices or functions [19].

An important instrument to investigate the structure of a vector space is the notion of linear independence. The vectors $v_1, \ldots, v_n$ are linearly independent, if all nontrivial combinations of the vectors are nonzero, i.e., $c_1v_1 + \ldots + c_nv_n \neq 0$ unless $c_1 = \ldots = c_n = 0$. For example, in three-dimensional space three vectors are dependent if they lie in the same plane. Four vectors are always linearly dependent in $\mathbf{R}^3$. If a vector space consists of all linear combinations of the specific vectors $w_1, \ldots, w_n$, then these vectors span the space. A basis for a vector space is a set of vectors, which is both linearly independent and spans the space.

In n-dimensional vector space, the length of a vector can be calculated by *n-1* applications of the Pythagoras formula, adding one more dimension at each step. The length of a vector in n dimensions is therefore $|v|^2 = v_1^2 + v_2^2 + \ldots + v_n^2$. The inner product of two vectors is used to test them for orthogonality. It is zero if and only if the vectors are orthogonal: $v * w = v_1w_1 + v_2w_2 + \ldots + v_nw_n = 0$.

2.2 Standardization of Variables

In order to calculate distances between instances of concepts in a conceptual space it is necessary that all variables (i.e., quality dimensions) of the space are represented in the same relative unit of measurement. In addition, the relations between elements regarding the values of a variable need to be obtained. A method, which fulfills these requirements, is calculating the z scores for these values, also called z-transformation [3].

The z score of a particular observation in a data set is

$$z_i = \frac{x_i - \bar{x}}{s_x}$$

where z_i is the i-th value of the new variable Z, x_i is the i-th value of the old variable X, $\bar{x}$ is the mean of X, and s_x is the standard deviation of X. The standard deviation therefore functions as the unit of measurement for describing the distances from the mean. It is positive or negative according to whether the original value lies above or below the mean. This transformation is also referred to as standardization [2].

[5] The full list of rules can be found in Strang [19].

3 Formalization of a Conceptual Space

In this section we describe how conceptual spaces can be formally represented as vector spaces. This allows for measuring semantic distances between instances of concepts and assigning weights to the dimensions of a space. Furthermore it is possible to define mappings between conceptual spaces.

3.1 Conceptual Vector Space

A conceptual vector space is defined as a set of vectors, each of which represents a particular quality dimension in the conceptual space. Ideally, these vectors form a basis of the space. In practice this is hard to achieve because for various domains not all dimensions are totally independent. For example, in the color domain, saturation and brightness influence each other [1]. If multi-domain concepts are represented in a conceptual space, then each quality dimension of the space can represent a whole domain, consisting of its quality dimensions. This way the conceptual vector space consists of a number of subspaces (and possibly subspaces of these, etc.).

Formally, a conceptual vector space is defined as $\mathbf{C}^n = \{(c_1, c_2, \ldots, c_n) \mid c_i \in \mathbf{C}\}$ where the c_i are the quality dimensions. If a quality dimension represents a domain, then the particular dimension $c_j = \mathbf{D}^n = \{(d_1, d_2, \ldots, d_n) \mid d_k \in \mathbf{D}\}$ and so on. One can best visualize such hierarchical structure as a rooted tree [21] G = (V, E) with the distinguished vertex representing the conceptual space and all other vertices representing either single quality dimensions or domains as sets of quality dimensions.

3.2 Distances and Weights

The fact that vector spaces have a metric allows for the calculation of distances between points in the space. These points are specific instances of concepts represented as vectors $v_{inst} = (v_1, v_2, \ldots, v_n)$ where the v_1 to v_n are the components, i.e., the values for the quality dimensions of the conceptual space. We refer to the Euclidean distances between points as *semantic distances* between instances of the concept represented in the particular conceptual vector space. Calculating the semantic distance between two instances of a concept u and v involves the following steps:

1. Calculation of z scores for components to get same relative unit of measurement:
   $(u_1, u_2, \ldots, u_n) \rightarrow (z_1^u, z_2^u, \ldots, z_n^u)$, $(v_1, v_2, \ldots, v_n) \rightarrow (z_1^v, z_2^v, \ldots, z_n^v)$
2. Calculation of semantic distance between u and v:
   $|d_{uv}|^2 = (z_1^v - z_1^u)^2 + (z_2^v - z_2^u)^2 + \ldots + (z_n^v - z_n^u)^2$

The definition of multiplication of vectors by real numbers in vector spaces makes it possible to assign weights to the individual quality dimensions. This is essential for the representation of concepts as dynamical systems [22]. Meanings of concepts change over time and depending on the context in which they are used. In a conceptual vector space it is possible to account for these changes by adding or deleting quality dimensions and by assigning different saliencies (as weights) to the existing dimensions. In this case $\mathbf{C}^n$ is defined as $\{(w_1c_1, w_2c_2, \ldots, w_nc_n) \mid c_i \in \mathbf{C}, w_j \in \mathbf{W}\}$ where $\mathbf{W}$ is the set of real numbers.

3.3 Mappings between Conceptual Vector Spaces

Conceptual vector spaces can be utilized for defining the matching of concepts between communication parties and to translate the meanings of concepts between different information communities [23]. This way they support the implementation of semantic reference systems [24, 25]. In order to do so it is necessary to define mappings between the spaces. Such mappings can either be transformations involving a major change in quality dimensions, e.g., addition of new dimensions, or projections reducing the complexity of a space by reducing its number of dimensions. Most mappings will lead to a loss of information.

Formally, such partial mappings are defined through mapping functions between components ($R: C^m \rightarrow C^n$). The impact of this definition is that for any given component $c^m \in C^m$, either there are no pairs $(c^m, c^n) \in R$ or there is only one such pair [21].

4 Case Study: Wayfinding Service with Landmarks

In this section we apply the formal methods to the case of a wayfinding service. This service offers facades of buildings as landmarks. The concept 'facade' is represented in a conceptual vector space, both from a system and a user perspective. It is further shown how the assignment of weights for the quality dimensions of 'facade' can be used to represent different contexts and how a simple mapping between system and user spaces can be performed.

4.1 Scenario

A wayfinding service communicating route instructions to its users serves as the case study to demonstrate the applicability and usefulness of conceptual vector spaces in a real-world scenario. The service automatically extracts salient features from datasets und uses these landmarks to enrich wayfinding instructions [4, 26]. The model of landmark saliency is based on three categories of attraction measures—visual, semantic, and structural. A particular wayfinding task in the city of Vienna had been used to apply the model and derive wayfinding instructions. Subsequent work included human subject tests to prove its cognitive plausibility [27].

The model has been applied to a limited case, i.e., pedestrians in a dense urban environment traveling in day-light and using facades of buildings as landmarks. The concept of 'facade' was represented by different variables, such as area, shape, and color. This represents essentially the system view. By extracting facades as landmarks based on these qualities and including them in wayfinding instructions for users, it is assumed that the system's concept of 'facade' equals the individual user's concept of 'facade'. This is most often not true. In the following we demonstrate how conceptual vector spaces can capture the differences between a system's view and a user's view of the same concept.

4.2 Conceptual Vector Space for 'Facade'

In order to represent the concept 'facade' in the conceptual spaces of both the system and the user, one first needs to define their quality dimensions. For the system view we take the original variables as used in [26]: facade area, shape factor, shape deviation, facade color on the RGB scale, visibility, cultural importance, and identifiability by signs. The dimensions for the user view are changed in two ways: First, it has been demonstrated that the metric of the RGB color space is usually different from the perceived metric [28], therefore a color model based on perception—the HSB (Hue, Saturation, Brightness) model, also called Natural Color System (NCS)—is used. Second, cultural importance is not taken into account as a quality dimension for the user concept because the fact that a building was designed by a famous architect is relevant for monumental protection but usually not recognizable by the average user of a wayfinding service.

Formally, the conceptual vector spaces for ‘facade’ can then be defined as $\mathbf{C}^n_{system}$ and $\mathbf{C}^n_{user}$:

- $\mathbf{C}^7_{system} = \{(c_1, c_2, \ldots, c_7) \mid c_i \in \mathbf{C}\}$ where
  c_1-c_3 correspond to the quality dimensions facade area, shape factor, and shape deviation;
  c_4 represents the color domain with RGB metric—$c_4 = \mathbf{D}^3 = \{(d_1, d_2, d_3) \mid d_i \in \mathbf{D}\}$ with d_1-d_3 being the quality dimensions red, green, and blue;
  c_5-c_7 correspond to the quality dimensions visibility, cultural importance, and identifiability by signs;
- $\mathbf{C}^6_{user} = \{(c_1, c_2, \ldots, c_6) \mid c_i \in \mathbf{C}\}$ where
  c_1-c_3 correspond to the quality dimensions facade area, shape factor, and shape deviation;
  c_4 represents the color domain with HSB metric—$c_4 = \mathbf{D}^3 = \{(d_1, d_2, d_3) \mid d_i \in \mathbf{D}\}$ with d_1-d_3 being the quality dimensions hue, saturation, and brightness;
  c_5-c_6 correspond to the quality dimensions visibility and identifiability by signs;

We can now represent individual facades, i.e., instances of the concept, as points[6] in these conceptual vector spaces. For illustration purposes we present these instances for a particular moment in the wayfinding task, i.e., the intersection Graben / Dorotheergasse in Vienna (Figure 1). Table 1 gives the z scores for the quality dimensions of the system space and table 2 shows the values for the user space. The original measured values for the quality dimensions were taken from [26]. Standardization was necessary because of the different units, such as square meters or percentages.

Figure 1: A 360° view of the intersection Graben / Dorotheergasse in Vienna.

In addition, the tables represent a point with id 0, which is an approximation to the prototypical facade region. This prototype is represented through the mean values of all quality dimensions for all facades considered. It is based on Rosch’s structural theory of centrality, where prototypical members had been found to correspond to the means of attributes that have a metric [7].

Making the simplifying assumption that the quality dimensions are independent, it is now possible to measure semantic distances between all instances and the prototype. This way, the conceptual vector space can be used for identifying the most distinct facade, i.e., the one with the largest distance from the prototype. It is exactly this facade, which should

[6] To be more precise: as vectors.

be used as a landmark for the wayfinding instructions. By looking at the distances and ranking of most distinctive facades one can notice that the selection of particular facades as landmarks for a specific user depends heavily on the concept of 'facade'. From the system's point of view facade 1 is the most distinctive and therefore chosen as the landmark at this intersection. From the user's perspective facade 2 is the best landmark. One of the reasons besides the differences regarding the color scale is the large semantic distance between facade 1 and the prototype with regard to cultural importance in the system space. In general, one can observe that the total ranking of facades differs widely between the two spaces.

Table 1: Instances of 'facade' from system's view at intersection Graben / Dorotheergasse.

id	z_1^c	z_2^c	z_3^c	z_1^d	z_2^d	z_3^d	z_5^c	z_6^c	z_7^c	dist	rank
1	-0.98	1.67	-0.38	-0.95	-0.84	-1.05	0.06	2.29	0.96	5.97	1
2	-1.09	0.76	-0.38	0.53	0.61	0.76	1.19	-0.64	0.96	5.40	2
3	0.33	-1.32	-0.38	-0.85	-0.79	-0.64	0.93	-0.64	0.96	4.25	6
4	-0.26	-0.02	-0.38	0.40	0.34	0.26	-2.08	-0.64	-1.18	4.62	5
5	0.53	-0.92	-0.38	2.02	2.09	2.09	-0.21	-0.64	0.96	5.28	4
6	-0.19	-0.18	-0.38	0.27	-0.02	-0.20	-0.58	-0.64	-1.18	4.15	7
7	-0.57	1.01	-0.38	-1.32	-1.34	-1.19	0.96	0.83	-1.18	5.33	3
0	2.24	-0.99	2.65	-0.10	-0.05	-0.03	-0.28	0.09	-0.32	0.00	-

Table 2: Instances of 'facade' from user's view at intersection Graben / Dorotheergasse.

id	z_1^c	z_2^c	z_3^c	z_1^d	z_2^d	z_3^d	z_5^c	z_6^c	dist	rank
1	-0.98	1.67	-0.38	-0.02	-0.33	-0.61	0.06	0.96	6.72	2
2	-1.09	0.76	-0.38	0.19	-0.37	0.71	1.19	0.96	7.12	1
3	0.33	-1.32	-0.38	0.18	-0.25	-0.34	0.93	0.96	5.84	7
4	-0.26	-0.02	-0.38	0.45	-0.44	0.32	-2.08	-1.18	6.62	5
5	0.53	-0.92	-0.38	0.24	-0.47	1.70	-0.21	0.96	6.66	3
6	-0.19	-0.18	-0.38	1.22	-0.46	-0.01	-0.58	-1.18	6.61	6
7	-0.57	1.01	-0.38	0.24	-0.32	0.21	0.96	-1.18	6.65	4
0	2.24	-0.99	2.65	-2.48	2.64	-1.99	-0.28	-0.32	0.00	-

4.3 Representing Different Contexts

People adapt their concepts to different decision situations. Prototypical regions in a conceptual space can change depending on the context and task at hand. With respect to wayfinding it has been argued that people alter their behavior due to variations of context, such as mode of traveling, role of the traveler, and environmental conditions: people focalize because they are offered different affordances in different decision situations and environments [29-31]. Winter et al. [32] demonstrated that people select different landmarks during wayfinding by day and night. Such variation of behavior can be modeled by assigning weights to the quality dimensions of a conceptual vector space.

The following formal definitions of two conceptual vector spaces represent the user's concept of 'facade' in the context of wayfinding by day and by night:

- $\mathbf{C}^6_{user\text{-}day} = \{(w_{1d}c_1, w_{2d}c_2, \ldots, w_{6d}c_6) \mid c_i \in \mathbf{C}, w_{jd} \in \mathbf{W}\}$
- $\mathbf{C}^6_{user\text{-}night} = \{(w_{1n}c_1, w_{2n}c_2, \ldots, w_{6n}c_6) \mid c_i \in \mathbf{C}, w_{jd} \in \mathbf{W}\}$

The weights w_{id} and w_{in} are assigned to the i-th quality dimensions for the day case and the night case. For calculation of the conceptual vector spaces these weights were taken from

[32][7] and slightly modified to fit the quality dimensions used here, e.g., the weight for shape was split into subweights for shape factor and shape deviation. The weights are shown in table 3.

Table 3: Weights for quality dimensions at day and night.

	z_1^c	z_2^c	z_3^c	z_1^d	z_2^d	z_3^d	z_5^c	z_6^c	Σ
w_{id} (day)	0.11	0.08	0.07	0.12	0.12	0.12	0.26	0.12	1.00
w_{in} (night)	0.26	0	0	0.07	0.07	0.07	0.23	0.30	1.00

Table 4 gives the new values for the semantic distances and the ranking of facades for the intersection Graben / Dorotheergasse. These results are derived from the user's conceptual vector space with the assigned weights for wayfinding at day and night.

Table 4: Semantic distances and ranking of facades for intersection Graben / Dorotheergasse at day and night.

id	$\mathbf{dist_{day}}$	$\mathbf{rank_{day}}$	$\mathbf{dist_{night}}$	$\mathbf{rank_{night}}$
1	0.70	6	0.97	2
2	0.84	1	1.06	1
3	0.69	7	0.75	6
4	0.83	2	0.88	4
5	0.74	4	0.71	7
6	0.73	5	0.78	5
7	0.76	3	0.89	3

One can observe the differences in values and the difference in ranking. Although facade 2 is chosen as the best landmark in both contexts, the distance from the prototype is 26% larger for the context of night. One of the reasons is the already large distance for the quality dimension area, which gets more than doubled through the high weight at night. Facade 1, which was ranked only sixth during daylight gets moved up to rank two at night. This is due to much higher gains for the now higher weighted dimensions (area and identifiability by signs) than losses for the lower weighted dimensions (color and shape).

4.4 Mapping from System to User Space

In order to bridge the semantic gap between the system's concepts and the user's concepts it is necessary to define mappings between their conceptual spaces. This is illustrated by a simple projection from the system's conceptual vector space for 'facade' to the user's conceptual vector space for 'facade' as defined in section 4.2. In this case the space gets reduced by one quality dimension (cultural importance) and the domain of color gets transformed from RGB to HSB. Formally, we can define the partial mapping (R: $\mathbf{C}^7_{system} \rightarrow \mathbf{C}^6_{user}$) with the following mapping functions between components: (c_1^s, c_1^u), (c_2^s, c_2^u), (c_3^s, c_3^u), $\{(d_1^s, d_1^u), (d_2^s, d_2^u), (d_3^s, d_3^u)\}$, (c_5^s, c_5^u), (c_7^s, c_6^u). For the color transformation rules an established algorithm[8] using color conversion formulas can be used. The use of such a mapping in our case study will ensure that the wayfinding instructions delivered by the ser-

[7] Weights were calculated from subjects' scoring of facades through a regression approach with a robust estimator.

[8] See, for example, http://webtools.arisesoft.com/colorset/.

vice will include landmarks according to the user's conceptualizations of ‘facade’. It is important to notice though that such transformations usually lead to losses of information or cannot be performed at all, i.e., when the necessary transformation rules cannot be identified or defined.

5 Conclusions and Future Work

This paper makes a contribution to formal representations of cognitive semantics. It describes a methodology to formalize conceptual spaces [1], which are sets of quality dimensions with a geometrical structure. Such spaces can be utilized for knowledge representation and sharing, and support the paradigm that concepts are dynamical systems. Based on the theory of vector space and a statistical standardization method, we demonstrated how individual conceptual vector spaces can be formally represented and concept instances compared by measuring semantic distances to a prototype. A wayfinding service using the concept of ‘facade’ and real-world data were used as a case study. Furthermore, it was shown how individual quality dimensions of a conceptual space can be assigned weights to account for different contexts. In order to bring the system’s semantics closer to the user’s semantics, conceptual vector spaces can be mapped from one to another in the way of transformations and projections.

The work leads to many directions for future research:

1. We made the simplifying assumption that the quality dimensions of the conceptual spaces are independent. This is not always true. It will be necessary to investigate the covariances between dimensions and to account for these in the representations of the conceptual spaces. Human subject tests are a possible way to identify the quality dimensions for a concept and to infer their dependencies, which would lead to non-orthogonal axes in the representation.
2. Prototypes are usually regions in the conceptual space. In this paper, the regions were approximated through points denoting the means of values for each dimension. Further work is necessary to identify the prototypical regions, which might best be represented by fuzzy boundaries.
3. Mappings between conceptual vector spaces were defined as partial mappings. As a consequence, for every dimension in the source space there is either exactly one or no dimension in the target space. There are cases though, where quality dimensions of a concept are split into two or more dimensions, i.e., they are refined, which essentially cannot be represented through functions. A possible approach into this direction is information flow theory [33].
4. Several researchers have argued against a geometric approach for concept representation and similarity measurement for the reasons that the axioms of minimality, symmetry, and triangle inequality do not hold cognitively. With conceptual vector spaces it seems possible to account for these phenomena by assigning different weights depending on the context. In this way, dissimilarity of the same concept depending on the viewpoint and asymmetric semantic distances could be represented. The axiom of triangle inequality seems to be violated only when different contexts are mixed (e.g., geographical and political), such as in the example given by Tversky [34].

Acknowledgments

I am grateful to Andrew Frank, Werner Kuhn, Michael Lutz, Clemens Nothegger, and Andreas Wytzisk for their input at various stages of this work. The comments from two anonymous reviewers provided useful suggestions to improve the content of the paper.

References

[1] P. Gärdenfors, *Conceptual Spaces - The Geometry of Thought*. Cambridge, MA: Bradford Books, MIT Press, 2000.

[2] J. Devore and R. Peck, *Statistics - The Exploration and Analysis of Data*, 4th ed. Pacific Grove, CA: Duxbury, 2001.

[3] G. Bahrenberg, E. Giese, and J. Nipper, *Statistische Methoden in der Geographie 1*, 2nd ed. Stuttgart - Leipzig: Teubner, 1999.

[4] M. Raubal and S. Winter, "Enriching Wayfinding Instructions with Local Landmarks," in *Geographic Information Science - Second International Conference, GIScience 2002, Boulder, CO, USA, September 2002*, vol. 2478, *Lecture Notes in Computer Science*, M. Egenhofer and D. Mark, Eds. Berlin: Springer, 2002, pp. 243-259.

[5] G. Lakoff, *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind*. Chicago: The University of Chicago Press, 1987.

[6] M. Knauff, C. Schlieder, and C. Freksa, "Spatial Cognition: From Rat-Research to Multifunctional Spatial Assistance Systems," *Künstliche Intelligenz*, pp. 5-9, 2002.

[7] E. Rosch, "Principles of Categorization," in *Cognition and Categorization*, E. Rosch and B. Lloyd, Eds. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1978, pp. 27-48.

[8] R. Green, "Internally-Structured Conceptual Models in Cognitive Semantics," in *The Semantics of Relationships - An Interdisciplinary Perspective*, R. Green, C. Bean, and S. Myaeng, Eds. Dordrecht, The Netherlands: Kluwer, 2002, pp. 73-89.

[9] M. Johnson, *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*. Chicago: The University of Chicago Press, 1987.

[10] G. Lakoff and M. Johnson, *Metaphors We Live By*. Chicago: The University of Chicago Press, 1980.

[11] G. Fouconnier and M. Turner, "Conceptual integration networks," *Cognitive Science*, vol. 22, pp. 133-187, 1998.

[12] W. Kuhn, "Modeling the Semantics of Geographic Categories through Conceptual Integration," in *Geographic Information Science - Second International Conference, GIScience 2002, Boulder, CO, USA, September 2002*, vol. 2478, *Lecture Notes in Computer Science*, M. Egenhofer and D. Mark, Eds. Berlin: Springer, 2002, pp. 108-118.

[13] G. Fouconnier, *Mappings in Thought and Language*. Cambridge: Cambridge University Press, 1997.

[14] K. Holmqvist, "Conceptual engineering: implementing cognitive semantics," in *Cognitive Semantics: Meaning and Cognition*, J. Allwood and P. Gärdenfors, Eds. Amsterdam: John Benjamins, 1999, pp. 153-171.

[15] A. Frank and M. Raubal, "Formal Specifications of Image Schemata - A Step to Interoperability in Geographic Information Systems," *Spatial Cognition and Computation*, vol. 1, pp. 67-101, 1999.

[16] A. Rodríguez and M. Egenhofer, "Image-Schemata-Based Spatial Inferences: The Container-Surface Algebra," in *Spatial Information Theory—A Theoretical Basis for GIS, International Conference COSIT '97, Laurel Highlands, PA*, vol. 1329, *Lecture Notes in Computer Science*, S. Hirtle and A. Frank, Eds. Berlin: Springer, 1997, pp. 35-52.

[17] W. Kuhn and A. Frank, "A Formalization of Metaphors and Image-Schemas in User Interfaces," in *Cognitive and Linguistic Aspects of Geographic Space*, *Behavioural and Social Sciences - Vol. 63*, D. Mark and A. Frank, Eds., NATO ASI Series ed. Dordrecht, Boston, London: Kluwer Academic Publishers, 1991, pp. 419-434.

[18] J. Aisbett and G. Gibbon, "A general formulation of conceptual spaces as a meso level representation," *Artificial Intelligence*, pp. 189-232, 2001.

[19] G. Strang, *Linear Algebra and Its Applications*, Second edition ed. Orlando, FL: Academic Press, 1980.

[20] A. Beutelspacher, *Lineare Algebra*, 6th ed. Wiesbaden, Germany: Vieweg, 2003.

[21] M. Piff, *Discrete Mathematics - An introduction for software engineers*. Cambridge: Cambridge University Press, 1991.

[22] L. Barsalou, "Situated simulation in the human conceptual system," *Language and Cognitive Processes*, vol. 5, pp. 513-562, 2003.

[23] OGC, "The OpenGIS® Guide," The Open GIS Consortium, 1998.

[24] W. Kuhn, "Semantic Reference Systems," *International Journal of Geographical Information Science*, vol. 17, pp. 405-409, 2003.

[25] W. Kuhn and M. Raubal, "Implementing Semantic Reference Systems," in *AGILE 2003 - 6th AGILE Conference on Geographic Information Science*, *Collection des sciences appliquees de l' INSA de Lyon*, M. Gould, R. Laurini, and S. Coulondre, Eds. Lyon, France: Presses Polytechniques et Universitaires Romandes, 2003, pp. 63-72.

[26] C. Nothegger, "Automatic Selection of Landmarks," in *Institute for Geoinformation*. Vienna: Vienna University of Technology, 2003, pp. 80.

[27] C. Nothegger, S. Winter, and M. Raubal, "Selection of Salient Features for Route Directions," *Spatial Cognition and Computation*, vol. 4, forthcoming 2004.

[28] H. Mallot and J. Allen, *Computational vision*. Cambridge: MIT Press, 2000.

[29] J. Gibson, "The Theory of Affordances," in *Perceiving, Acting, and Knowing - Toward an Ecological Psychology*, R. Shaw and J. Bransford, Eds. Hillsdale, New Jersey: Lawrence Erlbaum Ass., 1977, pp. 67-82.

[30] M. Raubal and M. Worboys, "A Formal Model of the Process of Wayfinding in Built Environments," in *Spatial Information Theory - Cognitive and Computational Foundations of Geographic Information Science, International Conference COSIT '99, Stade, Germany*, vol. 1661, *Lecture Notes in Computer Science*, C. Freksa and D. Mark, Eds. Berlin: Springer, 1999, pp. 381-399.

[31] M. Raubal, "Ontology and epistemology for agent-based wayfinding simulation," *International Journal of Geographical Information Science*, vol. 15, pp. 653-665, 2001.

[32] S. Winter, M. Raubal, and C. Nothegger, "Focalizing Measures of Salience for Route Directions," in *Map-based mobile services – Theories, Methods and Implementations*, *Springer Geosciences*, A. Zipf, L. Meng, and T. Reichenbacher, Eds. Berlin: Springer, forthcoming 2004.

[33] Y. Kalfoglou and M. Schorlemmer, "IF-Map: an ontology mapping method based on information flow theory," *Journal on Data Semantics*, vol. 1, pp. 98-127, 2003.

[34] A. Tversky, "Features of Similarity," *Psychological Review*, vol. 84, pp. 327-352, 1977.