MATEC Web of Conferences **146**, 01004 (2018) https://doi.org/10.1051/matecconf/201814601004
Building Defects 2017

Detecting, modeling and quantifying differences in terms meaning, in BIM environment

Miroslav Sůra[1,*]

[1]CTU Prague, Faculty of Civil Engineering, Thakurova 7, 166 29 Prague, Czech Republic

Abstract. Professional tasks in BIM involve sharing information among many participants, and require the use of the most precise terms to meet the following conditions. First, the user (reader) must not hesitate over the meaning. Second, various users should understand them alike, or better – in the same way. Less vague expressions provide more reliable basis for decision analysis and decision making, and for a joint, seamless, unquestionable consensus. Formal definitions of terms are used for that purpose where necessary, e.g. for logical or computer processing, but other terms remain or need to remain without such precise formal definitions. The principles for comparing difference in understanding the meaning of a term by different subjects are examined. We specify principles for measuring and calculating difference in, and vagueness of, meaning. Then, variables and scaling methods are proposed to measure the difference in and vagueness of meaning. The presented measuring scales can be applied for example in the process of refinement, unifying or splitting BIM terms, or group consistency in understanding a BIM-oriented terms can be assessed using the concept of vagueness.

1 Introduction

1.1 Vagueness

Vagueness [1] is a common feature of many tasks not only in the area of building, with or without BIM, and the tasks include communication. The vagueness is due to using natural languages. Building tasks as well as many other technical and professional tasks are often starting their existence from formulations in a natural language, too. Later on, however, the analysts strive to express the ideas in such a form that eliminates the vagueness. To this end, formal languages are used.

Some disciplines, like mathematics and physics, are completely relying on formal languages. Applied, and technical, disciplines also use formal tools to eliminate vagueness—to be effective and practical, to reach as exact and reliable results as possible. For some parts of the task, however, it is necessary to use natural language as a tool. This stems from the fact that these parts are not purely technical [2, 3] but take into account

* Corresponding author: sura@fsv.cvut.cz

human beings, social questions, emotional aspects, artistic views and so on [4]—and terms in descriptions using natural language are naturally vague.

In BIM environment, another additional feature applies besides the above—various professions come into contact via BIM, professions that otherwise might have less or no need to interact.

1.2 Semantic differential

The tasks in building area using BIM—or some of the sub-tasks—then combine 'formal descriptions' and 'natural (vague) descriptions'. The part with formal descriptions uses variables to express values, and ideally eliminates vagueness. Such a variable, that does not contain vagueness, has 'zero semantic differential'. Other variables, typically from the part with natural (and vague) descriptions, do not ensure the elimination of vagueness effectively; their semantic differential is non-zero, the participants need not to understand it in the same way.

The vagueness (of the meaning) of a variable need not always be a problem. Vagueness *is not* inaccuracy, incorrectness or unfaithfulness of the actual value of the variable, but vagueness also includes the awareness of the fact that other participant(s) may understand the terms in various different ways. Many practical tasks are satisfactorily tackled without strict descriptions and definitions of each relevant variable. In this case, removal of vagueness provides no substantial additional benefits. Typical case is if the solution exhibits only minor sensitivity to that variable change.

In other cases, however, the vagueness is a problem—e.g. when results are not interpreted, explained or assessed equally by every BIM participant and/or user. This may be due to semantic differential in some (both input and output) variable(s). Then locating and decreasing the semantic differential(s) is an adequate option as it reduces the danger of defects in any respective building life cycle phase (planning, construction, use…).

1.3 The case for semantic differential and vagueness

Some building associated terms like 'comfort of living' or 'road safety' are vague by themselves—in this case it is not a drawback but rather a merit and advantage. This derives from the fact that building isn't a static discipline but an evolving one, responding to new demands, scientific progress and technological innovations [5].

The content/meaning of such terms as 'comfort of living' changes in time [6]. These changes make the term more or less vague over time—and semantic differential varies over time, too. Some constituents of comfort are well defined, exhibit (almost) no vagueness, and (almost) no semantic differential (heat, humidity), others probably change (space requirements), others may be vague due to individual views, attitudes, location, culture, and also change over time (colors & lighting, furnishing, population density, amenities…) [7]. Not recognizing the related semantic differential change causes defects, in building plans, constructions…

2 Calculating semantic differential and vagueness

2.1 Dimensions, Aspect Ratings, Aspect Rating Density Function

Semantic differential is a type of a rating scale designed to measure the connotative meaning of objects, events, and concepts; and the semantics or meaning of words, particularly adjectives or featuring descriptors, and their referent concepts [8]. This is applicable also to building and BIM to encompass (some of the) variables defined for a building task. The connotative meaning is represented by a point in a space which is a Cartesian product of m *aspects* X_i (or *dimensions*) $G \equiv X_1 \times X_2 \times \ldots \times X_m$. If an individual aspect is a discrete scale, it is mapped on some suitable matching continuous scale. The aspects are defined using a pair of contrasting names (good–bad, slow–fast, elaborate–simplified) or a single name (desired population density, maximum walking distance, amenities availability, collision rate…). Viewed from the other side, the semantic differential approach helps in discovering and defining relevant variables if the problem is formulated in a natural language and a recognized or suspected vagueness is present.

A subject i rates the meaning separately for each aspect, which yields a point $\boldsymbol{a}^{(i)}$ within the *aspect space* G. The difference between two ratings $\boldsymbol{a}^{(i_1)}$ and $\boldsymbol{a}^{(i_2)}$ is the *semantic differential* in (the perception of) the meaning between the two subjects. It need not be simply a vector connecting the two points, but rather a calculated vector value:

$$\boldsymbol{d}^{(i_1\backslash i_2)} \equiv \langle d_{\boldsymbol{1}}^{(i_1\backslash i_2)},\ d_{\boldsymbol{2}}^{(i_1\backslash i_2)}, \ldots,\ d_{\boldsymbol{m}}^{(i_1\backslash i_2)}\rangle = \boldsymbol{u}\left(\boldsymbol{a}^{(i_1)},\ \boldsymbol{a}^{(i_2)}\right)\ . \qquad (2\ .1)$$

For ease of calculations, to make it more general and versatile, the ratings defined as a point $\boldsymbol{a}^{(i)}$ in G are extended to continuous *aspect rating density* function[†] or just *aspect rating* function $f^{(i)}(G)$.

We suppose that aspect rating function satisfies factorization condition (the dimensions are handled as mutually independent, rather uncorrelated, for each subject i of N subjects)

$$f^{(i)}(G) = f^{(i)}(x_1, x_2, \ldots, x_m) = \prod_{\boldsymbol{k=1}}^{\boldsymbol{m}} f_k^{(i)}(x_k)\ ,\quad i = 1, \ldots, N\ , \qquad (2\ .2)$$

and aspect rating-completeness or rating-unit-measure condition, and is non-negative

$$\int_{X_k} f_k^{(i)}(x_k)dx_k \quad = 1\ ,\qquad \forall\, x_k : f_k^{(i)}(x_k) \geq 0\ ,\qquad k = 1, \ldots m\ . \qquad (2\ .3)$$

To compare two aspect densities we define *semantic differential density* as two subjects' aspect densities difference (for $k = 1,2, \ldots, m$)

$$d_k^{(i_2\backslash i_1)}(x_k) = u_k^{(i_2\backslash i_1)}\left(f_k^{(i_2)}(x_k) - f_k^{(i_1)}(x_k)\right),\ \ \boldsymbol{d}^{(i_2\backslash i_1)} = \langle d_1^{(i_2\backslash i_1)}, \ldots, d_m^{(i_2\backslash i_1)}\rangle\ ,\quad (2\ .4)$$

where $\mathrm{u}_\mathrm{k}^{(\mathrm{i}_2\backslash \mathrm{i}_1)}$ may—but if appropriate need not—be an identity.

2.2 Aspects vagueness and semantic differential

Vagueness $vag_k^{(i)}$ of subject i's rating of k-aspect (or vagueness of subject i on aspect k) should depend on the rating of aspect x_k, i.e. on the aspect density function $f_k^{(i)}(x_k)$ and can be measured using statistical or probabilistic functions. Among the first to come into consideration is variance $var_k^{(i)}$, or optionally standard deviation $\sigma_k^{(i)}$, $var_k^{(i)} = {\sigma_k^{(i)}}^2$. Using variance for measuring vagueness $vag_k^{(i)}$ of k-aspect rating we define

$$vag_k^{(i)} \equiv var_k^{(i)} = \int_{X_k} \left(x_k - \mu_k^{(i)}\right)^2 f_k^{(i)}(x_k)dx_k\ ,\qquad k = 1, 2, \ldots, m\ . \qquad (2\ .5)$$

The mean $\mu_k^{(i)}$—or expected value E—of the aspect k rating function by subject i is

$$\mu_k^{(i)} = E\left[f_k^{(i)}(x_k)\right] = \int_{X_k} x_k\, f_k^{(i)}(x_k)dx_k\ ,\ \ k = 1, 2, \ldots, m\ . \qquad (2\ .6)$$

In vector form, vagueness and mean as related to subject i's rating are

[†] The conversion from discrete to a continuous scale can be performed using Dirac delta functions.

$$\boldsymbol{vag}^{(i)} = \langle vag_1^{(i)}, vag_2^{(i)}, \dots, vag_m^{(i)} \rangle . \qquad \boldsymbol{\mu}^{(i)} = \langle \mu_1^{(i)},\ \mu_2^{(i)}, \dots, \mu_m^{(i)} \rangle , \tag{2 .7}$$

Aspect density function mean being a counterpart of single-valued aspect rating on a discrete scale, the semantic differential can be viewed as a function based on (two) aspect rating densities means. If the function is the difference of the two means we have (in vector form)

$$\boldsymbol{d}_{\mu}^{(\boldsymbol{i_2} \backslash \boldsymbol{i_1})} = \boldsymbol{\mu}^{(\boldsymbol{i_2})} - \boldsymbol{\mu}^{(\boldsymbol{i_1})} = \boldsymbol{E}\big[\boldsymbol{d}^{(i_2 \backslash i_1)}(G)\big] \ . \tag{2 .8}$$

The above *semantic differential density* (2 .4) is 'directed'—it "points" from one subject's aspect rating to the other, and is or may be negative in some intervals. When trying to bring closer two or several participants in a building or BIM task, the 'directing' may turn to an inconvenience, it might increase aversions. To have an 'undirected' quantity ψ for semantic differential density, the following condition should hold

$$\psi\left(f_k^{(i_1)}(x_k)\ ,\ f_k^{(i_2)}(x_k)\right) = \psi\left(f_k^{(i_2)}(x_k)\ ,\ f_k^{(i_1)}(x_k)\right)\ \geq\ 0\ . \tag{2 .9}$$

One possible form of ψ is

$$\psi\left(f_k^{(i_1)}(x_k)\ ,\ f_k^{(i_2)}(x_k)\right) = \left|f_k^{(i_1)}(x_k) -\ f_k^{(i_2)}(x_k)\right|^n . \tag{2 .10}$$

The mean (2 .7) are vectors of numbers (one number per aspect) and thus the semantic differential (2 .8) cannot reflect the various courses (shapes) of aspect rating functions $f_{\mathrm{k}}^{(i)}(x_{\mathrm{k}})$. Other formulas for semantic differential can be used to reflect the 'shape' of aspect rating functions, for example

$$d_k^{(i_1 \top i_2)} = \sqrt{\int_{X_k} \left(x_k - E\left[exc_k^{(i_1 \top i_2)}(x_k)\right]\right)^2\ exc_k^{(i_1 \top i_2)}(x_k) dx_k}\ , \tag{2 .11}$$

$exc_k^{(i_1 \top i_2)}(x_k)$ may be defined by (2 .10) with $n = 1$, and $E\left[exc_k^{(i_1 \top i_2)}(x_k)\right]$ is its mean

$$exc_k^{(i_1 \top i_2)}(x_k) = \left|f_k^{(i_1)}(x_k) -\ f_k^{(i_2)}(x_k)\right|\ . \tag{2 .12}$$

3 Comprehension within an expert group

Mutual comprehension within a group of experts [9] extends the concept of *semantic differential* to a group of N subjects, each subject i within the group being assigned a weight $w_k^{(i)}$ for aspect k, and having his own aspect ratings $\mathbf{f}^{(i)}$, associated into *group ratings* $\mathbf{F}$, $\mathbf{F}_w$, and *group rating density functions* $\mathrm{f}_{\mathrm{k}}^{(\mathbf{F})}(\mathrm{x}_{\mathrm{k}})$, $\boldsymbol{f}^{(\mathbf{F})}$.

$$\boldsymbol{w} = \langle \boldsymbol{w}^{(1)}, \dots, \boldsymbol{w}^{(N)} \rangle ,\quad \boldsymbol{w}^{(i)} = \langle w_1^{(i)}, \dots,\ w_m^{(i)} \rangle\ ,\ \sum_{i=1}^{N} w_k^{(i)} = N,\ k = 1, \dots, m\ , \tag{3 .1}$$

$$\boldsymbol{f}^{(i)}(G) = \langle f_1^{(i)}(x_1), \dots,\ f_m^{(i)}(x_m) \rangle ,\quad \boldsymbol{F} = \langle \boldsymbol{f}^{(1)}(G), \dots,\ \boldsymbol{f}^{(N)}(G) \rangle ,\quad \boldsymbol{F}_w = \langle \boldsymbol{F}, \boldsymbol{w} \rangle\ , \tag{3 .2}$$

$$\boldsymbol{f}^{(\boldsymbol{F})} = \langle f_1^{(\boldsymbol{F})}(x_1), \dots, f_m^{(\boldsymbol{F})}(x_m) \rangle ,\quad f_k^{(\boldsymbol{F})}(x_k) = \frac{1}{N}\sum_{i=1}^{N} w_k^{(i)}\ f_k^{(i)}(x_k)\ ,\ k = 1, \dots, m\ . \tag{3 .3}$$

3.1 Group rating

Using the above, *weighted group mean rating* $\mu_k^{(\mathbf{F})} \equiv E_w^{(\mathbf{F})}[f_k(x_k)]$ on aspect k is defined as

$$E_w^{(\boldsymbol{F})}\left[f_k^{(\boldsymbol{F})}(x_k)\right] = \frac{1}{N}\sum_{i=1}^{N} w_k^{(i)}\, E\left[f_k^{(i)}(x_k)\right] = \frac{1}{N}\sum_{i=1}^{N} w_k^{(i)} \int_{X_k} x_k\, f_k^{(i)}(x_k)dx_k \quad , \tag{3 .4}$$

or in vector form:

$$\boldsymbol{E}(\boldsymbol{F_w}) = \langle E_w^{(\boldsymbol{F})}\left[f_1^{(\boldsymbol{F})}(x_1)\right],\ E_w^{(\boldsymbol{F})}\left[f_2^{(\boldsymbol{F})}(x_2)\right],\ \dots, E_w^{(\boldsymbol{F})}\left[f_m^{(\boldsymbol{F})}(x_m)\right]\rangle \ . \tag{3 .5}$$

Group semantic differential Gd may be defined as an extension of semantic differential of two subjects (2 .11) and may be defined in various ways, e.g. as an average of semantic differentials of all pairs of experts in the group

$$Gd_k^{(\boldsymbol{F})} = \frac{1}{N(N-1)}\sum_{i_1,i_2=1}^{i_1,i_2=N} d_k^{(i_1 \top i_2)}, \quad i_1 \neq i_2\,, \tag{3 .6}$$

or as an average of semantic differential of group mean rating density $\mathrm{f}_\mathrm{k}^{(\boldsymbol{F})}(\mathrm{x}_\mathrm{k})$ and each individual rating density $\mathrm{f}_\mathrm{k}^{(\mathrm{i})}(\mathrm{x}_\mathrm{k})$ (see (2 .11), (2 .12))

$$exc_k^{(\boldsymbol{F}\top i)}(x_k) = \left|f_k^{(\boldsymbol{F})}(x_k) - f_k^{(i)}(x_k)\right| \ , \tag{3 .7}$$

$$d_k^{(\boldsymbol{F}\top i)} = \sqrt{\int_{X_k} \left(x_k - E\left[exc_k^{(\boldsymbol{F}\top i)}(x_k)\right]\right)^2 exc_k^{(\boldsymbol{F}\top i)}(x_k)dx_k} \ , \tag{3 .8}$$

$$\widetilde{Gd}_k^{(\boldsymbol{F})} = \frac{1}{N}\sum_{i=1}^{N} d_k^{(\boldsymbol{F}\top i)} \ . \tag{3 .9}$$

In addition to mean and semantic differential we can define vagueness in aspect k in various ways similar to (2 .5)

$$GVag_k^{(\boldsymbol{F})} = \int_{X_k} \left(x_k - \mu_k^{(\boldsymbol{F})}\right)^2 f_k^{(\boldsymbol{F})}(x_k)dx_k = \int_{X_k} x_k^2\, f_k^{(\boldsymbol{F})}(x_k)dx_k - \left(\mu_k^{(\boldsymbol{F})}\right)^2 , \tag{3 .10}$$

$$\widetilde{GVag}_k^{(\boldsymbol{F})} = \frac{1}{N}\sum_{i=1}^{N} \int_{X_k} \left(x_k - \mu_k^{(i)}\right)^2 w_k^{(i)} f_k^{(i)}(x_k)dx_k \leq GVag_k^{(\boldsymbol{F})} \ . \tag{3 .11}$$

4 Conclusions

In collaborative building processes, mutual comprehension forms a substantial part of success. The concept of semantic differential offers a tool to identify potential or real threats of miscomprehension and to foresee and prevent defects in building processes as a consequence thereof. Data for semantic differential evaluation can be collected, or may be already available. For example, if an expert panel on some MCDA task evaluates the alternatives, the criteria play the role of aspects and each alternative is one of the objects rated on aspect scales. Calculating the vagueness and/or semantic differentials may provide hints of possible miscomprehensions.

- High semantic differential may be due to different opinions of experts, or to perceiving the underlying item in divergent ways. Or it may be due to different methods used by experts, or to novelty of the expert area and its development.
- Low semantic differential may be not due to equal comprehension, but to traditional views or to equal methods used by the participating experts. Or it may be due to lack of creative thinking.
- Finding a subgroup of experts exhibiting substantially lower semantic differential or vagueness may be a sign of some dissension, and unifying the views may be desirable. Or it may simply be a sign of several extant paradigms or schools.

Higher semantic differential or vagueness may indicate misunderstanding, although it may have other causes, too, and in the field of interest it may lead to discovering, besides miscomprehensions, also other features. It may also indicate the intention of some participants to keep the term vague ('social housing' being an example).

Funded as well as unfunded support from various European projects and foundations is gratefully acknowledged.

References

1. B.A.W. Russell, *Vagueness*, Australian Journal of Philosophy **1**, 84-92, (1923)
2. L.H. Dreyfus, *What Computers Still Can't Do: A Critique of Artificial Reason*. (MIT Press, New York, 1992)
3. R, Penrose, *The large, the small and the human mind*, (Cambridge University Press, 1997)
4. D. Kahneman, A. Tversky, *Prospect Theory*, Econometrica, **47(2)**, 263-291 (1979).
5. J.A. Laitner et al., *The Long-Term Energy Efficiency Potential: What the Evidence Suggests*. Available on http://amca.org/assets/member-document/6ACEEE-Vision-Report-2012.pdf (2012)
6. D. Madland, R. Teixeira, *New Progressive America: The Millennial Generation*. Available on https://cdn.americanprogress.org/wp-content/uploads/issues/2009/05/pdf/millennial_generation.pdf (2009)
7. J. Rifkin, *The zero marginal cost society: the internet of things, the collaborative commons, and the eclipse of capitalism* (Palgrave Macmillan, New York, 2014)
8. C.E. Osgood, G. Suci, P. Tannenbaum, The measurement of meaning (University of Illinois Press, Urbana, Illinois, 1957)
9. S. Plous, *The psychology of judgement and decision making* (McGraw-Hill, Inc., New York, 1993)