A Proposal of Quantification Method Describing the Difference between the Meaning of the Terms in the International Standards for Safety

Yohei Ueda
Department of Information Science and Engineering,
Shibaura Institute of Technology,
Tokyo Japan
Email: ma16013@shibaura-it.ac.jp

Masaomi Kimura
Department of Information Science and Engineering
Shibaura Institute of Technology,
Tokyo Japan
Email:masaomi@sic.shibaura-it.ac.jp

***Abstract*—International safety standards define and regularize many aspects of product safety during manufacturing processes. However, principles in international standards contain many homographic keywords or words with similar but slightly different meanings, which can cause ambiguity. We propose a method to quantify the differences in the meanings of keywords. We focus on the different meanings of definition statements, different dependency relationship structures, and different tendencies of the dependency relationships.**

Keywords–safety; homograph; international standard; semantics;

I. INTRODUCTION

Nowadays, human injury is caused by many daily-use products such as electronic devices, toys, and bicycles. This has recently become a critical situation that requires attention. Injuries are difficult to avoid by only "human attention." Therefore, product manufacturing processes must adhere to established international safety standards. The primary purpose of such standards is to define fundamental safety principles for product creation.

However, the statements of principles often include many homographic keywords, i.e., words that are spelled the same but have different meanings or words with similar but slightly different meanings; thus, their meaning may be ambiguous. Misinterpretation of the meanings of terms may cause difficulty to discuss in the International Organization for Standardization. Moreover, product designers may not adhere to the standards; this may result in manufacturers producing inappropriate products. Certification authorities might authenticate a dangerous product by mistake.

For example, the ISO/IEC Guide 51 safety standard, defines "risk" as a "combination of the probability of occurrence of harm and the severity of that harm." On the other hand, the ISO/IEC Guide73 risk management standard defines "risk" as the "effect of uncertainty on objectives." Clearly, the meaning of "risk" differs in these two standards.

In this study, we introduce a method to quantify the difference in meanings of safety terms based on international standards by using the content of terms and definitions and other elements (e.g., risk assessment and risk reduction). We focus on three types of differences, i.e., the meanings of definition statements, dependency relationship structures, and dependency relationship tendencies.

II. METHODS

A. Difference calculation in definition statements

In this study, we focused on essential details included in the "Terms and definitions" chapter of international standards. This chapter provides definitions of safety -related terms. Such definitions are very important for quantifying the difference between the meanings of terms in such documents.

We considered the definition statements that contain many important words for characterizing those statements. Because of the role of the "Terms and definitions" chapter, important words must be modified by other words to limit their meanings. In other words, a word with many dependency relationships can be regarded as more important in a sentence than words with fewer dependency relationships. For example, if we compare an "event" and a "harmful event," the latter meaning of "event" can be more stressed than the former; this represents a generic event.

Our method focuses on quantifying differences between the meanings of terms in definition statements. It measures the weights of words by estimating how meanings are limited for reducing ambiguity. In addition, the proposed method calculates a distance d_{def} between two international standards A and B.

To calculate this distance, we define a weight of word importance for a definition statement, v_{ids}.

According to our observations, most definition statements are noun phrases rather than sentences; this contributes to an incorrect dependency analysis. Therefore, to change the noun phrases to sentences, the phrase "this is" was added at the beginning of each definition statement.

We used the Stanford Parser to extract a parse tree based on dependency relationships in sentences. Degree centrality in the parse tree was used to identify the ratio of modifying or modified words to the total number of dependency relationships. Since the degree centrality should be high if a word has many dependency relationships, the word with a high degree centrality is considered as important. Therefore, v_{ids} can be expressed as follows:

$$v_{ids}(w) = \frac{k(w)}{\sum_{w \in W} k(w)}, \tag{1}$$

where $k(w)$ denotes the degree of the node representing the word w in a word set W in the sentence.

Note that from the viewpoint of importance of words, stop words (e.g., "a", "is", "the", and "from") can create noise; thus, stop words were not included in the word set during analyses. In addition, after parsing, the appended phrase "this is" was also not considered. We refer to the obtained words as "word groups."

After obtaining the words with their corresponding v_{ids} values in the word groups from international standards A and B, their distance (d_{def}) values were calculated on the basis

 ISBN: 978-1-61208-507-4

of the concept of Levenshtein distance. Levenshtein distance is a well-known measurement of the difference between two strings and is calculated as the minimum number of insertions and deletions of characters required to transform one string into the other. In this study, rather than characters, words in the word groups were replaced. The distance d_{def} is given as follows:

$$d_{def} = \sum_{w \in R} v_{ids}(w), \tag{2}$$

where R is a set of words added or deleted to make the word group for A coincide with that for B. Note that the words $w \in R$ only appear in one standard. Thus, we used the value of $v_{ids}(w)$ calculated in the network where w appeared.

B. Calculation of word meaning difference in body text

The above mentioned method does not cover the case in which words are used differently in the body text of international standards documents but their difference is not clarified in "Terms and definitions."

Therefore, we suggest two quantification methods using the body text of international standards documents. The first method is based on the structure of dependency relationships in the body text. The corresponding quantified value is expressed by $d_\phi = |\phi_A - \phi_B|$. The second method is based on latent semantics appearing in a dependency relationship tendency in the body text. Here, the quantified value is denoted d_{cos}.

1) Quantification of the difference in meaning using a dependency relationship structure: As discussed in Section II-A, the meaning of a word is limited when it is modified by other words, and this may cause a difference in meaning and importance. In other words, if there is a difference between the importance of the same word in different standards, there should also be a difference in meaning.

To extract this difference, we introduce an importance index $\phi(w)$ for a word w in the body text of an international standards document. The index value should be higher, if the word is limited and should be modified by words that are also important and have a high ϕ value. If $\phi_A(w)$ and $\phi_B(w)$ are the ϕ values of w for international standards A and B respectively, $|\phi_A(w) - \phi_B(w)|$ gives the difference in the extent of importance of the word w in the body text of the standards.

First, dependency analysis was applied to extract dependency relationships of the words in the body text of a given standard. Then, each verb was changed to its prototype and stop words were removed. Dependency relationships were represented as edges in a network in order to express the word relationships in the body text. Words in sentences were considered as nodes.

Secondly, after generating the networks, an importance value was assigned to each node using PageRank method. PageRank assigns a higher value to a node linked to many other nodes that have high values. In this study, the PageRank value corresponds to a higher importance value assigned to a word that modifies or is modified by the various and important words. In this paper, we denote the value assigned to word w by PageRank as $\phi(w)$.

Finally, the difference in meaning on the basis of the dependency structure was calculated as $|\phi_A(w) - \phi_B(w)|$. We used d_ϕ as a difference index for $|\phi_A(w) - \phi_B(w)|$.

2) Difference in meaning of the tendency of dependency relationship: Section 2.2.1 showed how to quantify the difference in meaning using the structure of modification. However, if two words are modified by the same number of words, the difference could be low. For example, international standards A and B have dependency networks around the word "train," as shown in Figs. 1 and 2. However, in international standard A, "train" means teaching someone, and in international standard B, "train" means a railway train.

In this case, $|\phi_A(train) - \phi_B(train)|$ could be low because they have the same structure. In this section, we show that the difference between word meanings in international standards A and B are quantified by expressing the frequency tendency of dependency relationships as a matrix, applying latent semantic analysis, and calculating the distance.

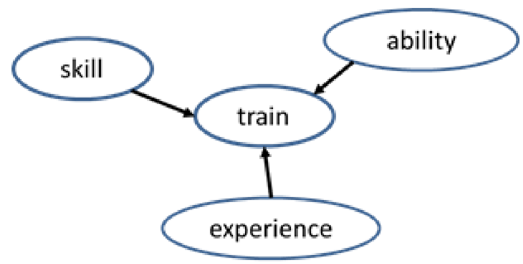

Fig. 1. International Standard A

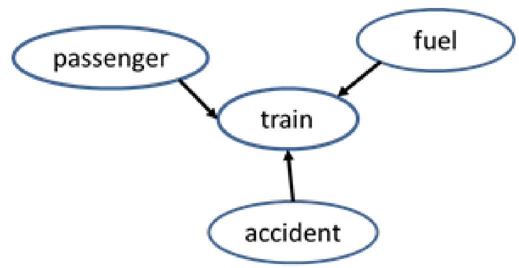

Fig. 2. International Standard B

First, dependency parsing was applied to the sentences in the body text of each standard. Second, a matrix was used to quantitatively express words that have dependency relationships with other words. The matrix M has frequencies of dependency relationships as elements. The rows correspond to modified words that commonly appeared in both standards, and the columns correspond to the modifying words in the standards that are handled separately. For example, if word w modifies word w' with frequency f_{w_A} in international standard A and with f_{w_B} in B, the matrix has a row w', columns w_A (denoting w in A) and w_B (denoting w in B), and elements f_{w_A} and f_{w_B} respectively.

It is convenient to use a vector space model to calculate the semantic similarity of words. To reduce the effect of noise, we employed latent semantic indexing(LSI) with singular- value decomposition. By applying singular- value decomposition, M can be decomposed as follows:

$$M = U \Sigma V^T, \tag{3}$$

where U and V are orthogonal matrices and Σ is a singular value matrix. The column vectors of V gives principal eigen-

 ISBN: 978-1-61208-507-4

vectors corresponding to the modifying words whose cosine values were used to measure similarity. Let $\vec{v}(w_A)$ and $\vec{v}(w_B)$ denote the principal eigenvectors for w_A and w_B respectively. Then, their cosine value is given as follows:

$$\cos\theta(\vec{v}(w_A), \vec{v}(w_B)) = \frac{\vec{v}(w_A)\cdot\vec{v}(w_B)}{|\vec{v}(w_A)||\vec{v}(w_B)|}, \quad (4)$$

whose value is in the range [0, 1]. However, in this study, we need to use distance rather than similarity. Therefore, we set the distance d_{cos} as follows:

$$d_{cos} = 1 - \cos\theta(\vec{v_a}, \vec{v_b}). \quad (5)$$

We used d_{cos} as a difference index for word meanings based on the tendency of the meaning of words.

C. Quantification of the difference in the meaning of a term

Finally, we merged the above mentioned indices, d_{def}, d_ϕ and d_{cos}. On the basis of the idea that the merged index needs to take a larger value if some or all of the indices have a large value, we designed it to be a linear combination as follows:

$$D = \alpha d_{def} + \beta d_{cos} + \gamma d_\phi, \quad (6)$$

where α,β and γ are coefficients to adjust the ranges of the indices.

To determine α, β and γ, we investigated the ranges of d_{def}, d_ϕ and d_{cos} by calculating them for words that appear in both the ISO/CD Guide51 and ISO12100 standards. We found that the ranges were completely different. Since the average values of d_{def}, d_ϕ and d_{cos} were 0.323, 0.41 and 0.0019, respectively, we set $\alpha = 1.28$, $\beta = 1$ and $\gamma = 250$ such that the product of the average values and the coefficient for each index equaled one. We expected that the coefficients would allow the ranges to be in the same order and confirmed that this method was effective by investigating other document pairs.

III. EXPERIMENTS

We conducted experiments to evaluate the D value by comparing it with each pair in four international standards. The standards used in our experiment. are listed in TABLE I. TABLE II lists the paired standards used in our experiment.

TABLE I. STANDARDS USED IN OUR EXPERIMENT

Standards	Summary
ISO/CD Guide 51	Safety aspects Guidelines for their inclusion in standards
ISO 12100	Safety of machinery
ISO/IEC Guide 50	Safety aspects Guidelines for child safety
ISO 8124	Safety of toys Age determination guidelines

TABLE II. PAIRED STANDARDS IN OUR EXPERIMENTS

	Standard A	Standard B
#1	ISO/CD Guide 51	ISO 12100
#2	ISO/IEC Guide 50	ISO 8124

We observed the D value containing d_{def}, d_{cos} and d_ϕ obtained by our method.

TABLE III. d_{def}, d_{cos}, d_ϕ AND D VALUES BETWEEN ISO/CD GUIDE 51 AND ISO 12100

Words	d_{def}	d_{cos}	d_ϕ	D
harm	0.575	0.502267	0.00010945	1.50713
standard	0	0.504292	0.01577331	4.44761
train	0	0.097685	0.00016153	0.13807

TABLE IV. d_{def}, d_{cos}, d_ϕ AND D VALUES BETWEEN ISO/IEC GUIDE 50 AND ISO 8124

Words	d_{def}	d_{cos}	d_ϕ	D
harm	0.25	0.262754	0.00134539	1.024101
edge	1	0.421634	0.00630093	3.276866
period	0	0.139401	0.00029886	0.214116

TABLE III and TABLE IV show the d_{def}, d_{cos}, d_ϕ and D values for each standard.

As shown in TABLE III, the D value for “standard” was greater than that for “harm” and “train.” Clearly, there was a large difference in the meaning of the word “standard” between ISO/CD Guide 51 and ISO 12100. The D value of “train” was smaller than that of “harm” and “standard.” The d_{def} value of “harm” was greater than that of “standard” and “train.” As shown in TABLE IV, the D value of “edge” was greater than that of “harm” and “period.” Clearly, there was a large difference in the meaning of the word “edge” between ISO/IEC Guide 50 and ISO 8124. The D value of the word “period” was smaller than that of “edge” and “harm,” and the d_{def} value of “harm” was greater than that of “period” and “edge.”

IV. DISCUSSION

A. Comparison of ISO/CD Guide51 and ISO12100

Here, we discuss our index values for the words “harm,”which is defined in “Terms and definitions” in each standard, with “standard” having a large D value and “train” having a small D value.

- **harm** The word “harm” had large d_{def} and d_{cos} values. Definition statements in both standards included the meaning “physical injury or damage to the health of people,” while definition statements in Guide51 included “damage to property or the environment.” This difference influenced the d_{def} value. TABLE V lists the words modified by “harm” and their frequency in ISO/CD Guide51 and ISO12100.

TABLE V. WORDS WITH A DEPENDENCY RELATIONSHIP WITH “HARM” AND THEIR FREQUENCY

	dependency relationship	Frequency
ISO/CD Guide51	harm → present	4
	harm → eliminate	1
	harm → avoid	4
ISO12100	harm → severity	3
	harm → occurrence	3

As can be seen in TABLE V, there was no common word modified by “harm” in ISO/CD Guide51 and ISO12100; this resulted in high d_{cos} values. No definition statements limited the meaning of “harm;” this resulted in a small ϕ and, therefore, a low value of d_ϕ. Thus, we observed that the values of the

 ISBN: 978-1-61208-507-4

indices included in D coincide with the situation related to “harm.”

- **standard** The large value of D for “standard” originates in the value of d_ϕ.
  TABLE VI shows the total number of words that modify “standard” in Guide51 and ISO12100. As can

TABLE VI. TOTAL NUMBER OF WORDS THAT MODIFY “STANDARD” IN GUIDE51 AND ISO12100

	Guide51	ISO12100
Total	22	3

be seen in Table VI, more words modified “standard” in Guide51 than in ISO12100, i.e., in Guide51, “standard” is limited more by other words than in ISO12100. We believe this is because Guide51 is an introductory safety guideline; thus, “standard” is modified by many other words. This means its modification structures give different ϕ values and high d_ϕ values.

- **train** The word “train,” which has a small D value, did not have a definition statement in either of the standards. Furthermore, each instance of “train” had few dependency relationships with other words. Regarding d_{cos}, there was a common dependency relationship, i.e., “train $\rightarrow$ skill,” between Guide51 and ISO12100. This shows that both standards use “train” to mean “teach.” In addition, “train” was not modified in either standard and, thereby, d_ϕ took small values; consequently, d_ϕ took a small value. Therefore, we obtained a small D value.

B. ISO/IEC Guide50 and ISO 8124

Words subject to evaluation were “harm,” which had definitions in “Terms and definitions” in each standard; “edge,” which had a large D value; and “period,” which had a small D value.

- **harm**
  The word “harm” had different meanings in definition statements, i.e., “physical injury” or “injury.” However, there was a common statement “damage to the health of people, or damage to property or the environment.” Therefore, we obtained a small d_{def} value. The value of d_ϕ was relatively high. TABLE VII shows the total number of words that modify “harm” in ISO/IEC Guide50 and ISO 8124.

TABLE VII. TOTAL NUMBER OF WORDS THAT MODIFY “HARM” IN ISO/IEC GUIDE50 AND ISO 8124

	Guide50	ISO 8124
Total	9	3

There were more words that modified “harm” in ISO/IEC Guide50 than in ISO 8124, i.e., “harm” in Guide50 is more limited by other words than in ISO 8124, which gave the high d_ϕ value.

- **edge**
  The word “edge” had a large d_ϕ value. TABLE VIII shows the total number of words that modify “edge” in ISO/IEC Guide50 and ISO 8124. There

TABLE VIII. TOTAL NUMBER OF WORDS THAT MODIFY “EDGE” IN ISO/IEC GUIDE50 AND ISO 8124

	ISO/IEC Guide50	ISO 8124
Total	5	47

are more words that modify “edge” in ISO8124 than in Guide50. In the body of Guide50, “edge” had an abstract meaning, e.g., “corner, end.” However, there are many sentences that define “edge” specified with concrete values in ISO8124. Furthermore, the value of d_{def} was 1 because although there was no definition statement for “edge” in Guide51, there was a definition statement for “edge” in ISO8124, i.e., “line, formed at the junction of two surfaces, whose length exceeds 2,0 mm.” We considered the case without a definition statement as a definition with “no word.” Moreover, we obtained a high d_ϕ value because of the difference in modification structures. Therefore, a large D value was obtained for “edge.”

- **period** The word “period” had a small D value, which was primarily because of the small d_ϕ value. TABLE IX shows the words that have a dependency relationship with “period” and their frequencies in ISO/IEC Guide50 and ISO 8124.

TABLE IX. WORDS THAT HAVE DEPENDENCY WITH “PERIOD” AND THEIR FREQUENCIES(ISO/IEC GUIDE50)

	dependency relationship	Frequency
ISO/IEC Guide50	certain $\rightarrow$ period	1
	extend $\rightarrow$ period	1
	long $\rightarrow$ period	1
	time $\rightarrow$ period	2
ISO 8124	h $\rightarrow$ period	1
	time $\rightarrow$ period	1

According to Table IX, the difference in the number of words that modify “period” was 2. There were more words that directly modify “period” in ISO/IEC Guide50. In contrast, the word “h,” which is a unit for “hour” in ISO 8124, was modified by many numbers, such as “72.” Therefore, there were many words that modify “period” indirectly.
Consequently, the value of d_ϕ was small because the difference of meaning based on limiting a meaning was not large.

V. CONCLUSION

Principles in international standards contain many homographic keywords that may cause ambiguity for readers. Furthermore, previous studies have not quantified the difference in the meanings of words in international standards because this is a difficult task to know it.

In this study, we proposed a method and indices to quantify this difference, on the basis of three types of differences the meanings of definition statements, structure of dependency relationships, and tendency of dependency relationships.

 ISBN: 978-1-61208-507-4

For the first difference, the source of the extent of difference was considered to be the relationships of word modifications in definition sentences, weighting, and calculation of distances. Based on this, we proposed the index d_{def}.

For the second difference, it was considered to the different structures of dependency relationships. The index d_ϕ was calculated by creating networks that express the relationships among words in the complete text of the standards and by applying PageRank to the networks.

For the third difference, the idea of a latent semantic index was applied to obtain the trends of word meanings. The value of d_{cos} was computed as the cosine similarity of the obtained characteristic vectors.

Finally, we combined these three indices to obtain an index D to evaluate different word meanings in international safety standards.

Consequently, a high D value was obtained when the number of words with different dependency relationships and the number of different meanings for the definition statements were large and vice versa.

REFERENCES

[1] M. Mukaidono,"The basic concept of safety design", Japanese Standards Association, Tokyo, 2007, ISBN: 4542404056

[2] International Organization for Standardization, COMMITTEE DRAFT ISO/CD Guide51, ISO/TC COPOLCO / SC IEC/ACOS N 49, 2012

[3] Y. Miyazaki; T. Shioi, and K. Hatano, "A Word Sense Disambiguation Method based on Dependency Relations", "Information Science and Technology Forum Proceedings", no. D-021, Aug. 2015, pp.111-112

[4] T. Michishita; K. Nakayama; T. Hara; S. Nishio, and "A Word Sense Disambiguation Method by Using Topics of Documents in Wikipedia," Information Processing Society of Japan Kansai Branch Conference Papers", no. C-12, 2009

[5] International Organization for Standardization, Safety of machinery General principles for design Risk assessment and risk reduction, ISO 12100:2010(E), 2010

[6] M. Ishida, Text mining introductory study in R, Morikita Publishing Co., Ltd., Tokyo, ISBN: 4627848412

[7] H. Toyoda, An introduction to Data mining, Tokyo Tosho Co., Ltd., Tokyo, ISBN: 4489020457

[8] T. Suzuki, Data science to learn in R 8 Network analysis, Kyoritsu Publishing Co., Ltd., Tokyo, ISBN: 4320019288

[9] The Stanford Natural Language Processing Group, The Stanford Parser: A statistical parser, http://nlp.stanford.edu/software/lexparser.shtmlhttp://nlp.stanford.edu/software/lex-parser.shtml/[accessed: 2014-01-12].

[10] International Organization for Standardization, Safety aspects - Guidelines for child safety in standards and other specifications, ISO/IEC Guide50:2014, 2014

[11] International Organization for Standardization, Safety of toys Part 1:Safety aspects related to mechanical and physical properties, ISO 8124-1:2009(E), 2009

[12] International Organization for Standardization, Cycles - Safety requirements for bicycles, ISO 4210:2014, 2014

[13] International Organization for Standardization, Bicycles for young children - Safety requirements and test methods, EN 14765:2005, 2005