

Techniques to Aggregate the Characteristics of Internal Quality of an IS Specification Language

Albertas ČAPLINSKAS, Jelena GASPEROVIČ

*Department of Software Engineering, Institute of Mathematics and Informatics
Goštauto 12, 01108 Vilnius, Lithuania
e-mail: alcapl@ktl.mii.lt, j.gasperovic@algoritmusistemas.lt*

Received: May 2005

Abstract. The paper analyses the problem of aggregation of internal quality characteristics of an information system specification language. It surveys main categories of aggregation operators, their mathematical and behavioural properties and proposes how to aggregate quality characteristics taking into account measurements of all lowest-level characteristics in a proper way. The paper shows that four different kinds of dependencies can occur among characteristics of internal quality and discusses in detail what aggregation techniques are reasonable for each case. It also describes a heuristic for minimisation of possible deviations of measurement results obtained after using a number of different metrics or because of the inaccuracy of used measurement techniques.

Key words: IS specification languages, internal quality, quality characteristics, aggregation methods.

1. Introduction

The ISO/IEC 9126 standard (ISO/IEC 9126, 1991), distinguishes three kinds of quality: internal quality, external quality and quality in use. Internal quality describes the level to which the product was developed following “good engineering” practices. In other words, it evaluates the product from the point of view of developers. External quality describes the level to which the product correctly provides the expected services. It treats the product as “black box” and evaluates it from the point of view of potential users. Quality in use describes the level to which a product used by particular users meets their needs to achieve some specified goals. External quality of software in many aspects depends on its internal quality. Sometimes it is not simple to distinguish external quality and internal quality at all. Although in the field of Software Engineering such distinction is deeply reasonable, because “good engineering” practices play in this field very important role, in many other fields, including specification languages, the situation is slightly different. To distinguish internal and external quality of a specification language is very difficult and sometimes even impossible. In addition, there are no serious reasons to make such distinction for specification languages. Thus, in this paper, we will distinguish only two kinds of quality: internal quality and quality in use and use the term “internal quality” to address both internal and external quality. We understand internal quality of a specification language as the descriptive characteristic of a language as a product independently

from any context of use. It means that internal quality is described by some value obtained from measurements. Meanwhile, quality in use is evaluative characteristic of a language obtained by making a judgment based on criteria that determine the worthiness of a language for a particular project. In previous works (Čaplinskas *et al.*, 2002; Čaplinskas and Gasperovic, 2005a) a quality model to evaluate quality in use and taxonomy of characteristics of internal quality have been proposed. The main purpose of this paper is to propose aggregation techniques required to calculate internal quality from the results of measurements. Aggregation techniques are used for several purposes: to aggregate different sub-characteristics of internal quality; to aggregate results of several measurements of a particular quality sub-characteristic obtained using the same metric; and to aggregate results of measurements of a particular quality sub-characteristic obtained using several different metrics.

In the first case the problem is to bring to light what kinds of dependencies can occur among quality sub-characteristics in the taxonomy of characteristics of internal quality and to decide in which way to aggregate sub-characteristics in order to take into account all measurements in a proper way. The paper is devoted mainly to the solution of this problem.

In last two cases the problem is to minimise possible deviations generated by shortcomings of the particular metric or by inaccuracy of the particular measurement. The paper considers this problem too but only in a very limited extent. It does not provide any theoretical analysis of the problem, however, it proposes some heuristic, which is acceptable in most of practical situations.

The rest of the paper is organised as follows. Section 2 briefly surveys the theory of aggregation. Section 3 describes possible kinds of aggregation operators. Section 4 contains the main results of the paper. It analyses what kinds of dependencies among quality sub-characteristics can occur in the taxonomy of characteristics of internal quality and for each situation proposes methods to aggregate appropriate sub-characteristics. Section 5 proposes a heuristic to minimise deviations generated by shortcomings of the particular metric or by inaccuracy of the particular measurement. Section 6 concludes the paper.

2. Preliminaries

Aggregation of information plays an important role in many fields. According to (Detyniecki, 2000) the purpose of aggregation is “the simultaneous use of different pieces of information (provided by several sources) in order to come to a conclusion or a decision”. More exactly, “Aggregation refers to the process of combining values (numerical or non numerical) into a single one, so that the final result of aggregation takes into account in a given fashion all the individual aggregated values” (Grabisch *et al.*, 1998). A number of approaches, including rule-based approach, neuronal networks, fusion specific techniques, probabilistic approach, evidence theory, possibility theory, fuzzy set approach, and many other approaches (e.g., data aggregation approach (Adomenas and Ciucelis, 2002)), have been proposed for this aim. However, all proposed approaches are based

at some extent on the numerical aggregation. In other words, all approaches include aggregation of some numerical values. More generally, “the aggregation operators are mathematical objects that have the function of reducing a set of numbers into a unique representative (or meaningful) number” (Detyniecki, 2000). In theoretical considerations it is assumed usually that both aggregated values and result of aggregation belong to some finite interval, say interval $[0,1]$, without any assumption about their nature. The choice of the interval $[0,1]$ is not restrictive because any interval can be transformed into this interval using a positive linear transformation $ax + b$, $a > 0$ (Grabisch *et al.*, 1998). Then an aggregation operator is defined as a function that assigns a real number y ($y \in [0, 1]$) to any n -tuple $(x_1, x_2, \dots, x_n \mid x_i \in [0, 1])$.

Fundamental properties of aggregation operators have been considered by (Mayor and Trillas, 1986; Ovchinnikov, 1998; Mesiar and Komorníková, 1997; Detyniecki, 2000). These properties can be divided into two groups: mathematical properties and behavioural properties. To be intuitively meaningful, an aggregation operator must satisfy at least the following axioms (Detyniecki, 2000):

$$\text{Aggreg}(x) = x \quad (\text{identity}), \quad (2.1)$$

$$\text{Aggreg}(0, \dots, 0) = 0, \quad \text{Aggreg}(1, \dots, 1) = 1 \quad (\text{boundary conditions}), \quad (2.2)$$

$$\text{Aggreg}(x_1, x_2, \dots, x_n) \leq \text{Aggreg}(y_1, y_2, \dots, y_n) \quad (\text{monotonicity}). \quad (2.3)$$

$$\text{if } (x_1, x_2, \dots, x_n) \leq (y_1, y_2, \dots, y_n).$$

These axioms define fundamental mathematical properties that are inherent to any aggregation operator.

It is proven (Detyniecki, 2000) that any Aggreg operator is continuous with respect to each variable, associative, commutative, bisymmetric, and idempotent. However, these properties are non-axiomatic and can be derived from the axioms (2.1), (2.2) and (2.3). Besides, it is proven that some specific classes of aggregation operators have such mathematical properties as a neutral element, an absorbent element (annihilator), compensation (Pareto property), counterbalancement and reinforcement. Finally, some aggregation operators can be stable with respect to linear changes of measurement scale and invariant for any bijection (e.g., in case when aggregation operator is defined as a projection).

Behavioural properties are specific for each class of aggregation operators. In other words, each class has individual behavioural properties. In general, there exist four behavioural properties:

- possibility to express weights of importance on the values that are aggregated,
- possibility to express relations between the values that are aggregated,
- possibility to take into account the aims of aggregation, and
- possibility of an easy semantic interpretation (i.e., being able to relate the values that are aggregated to the behaviour implied by aggregation operator).

3. Aggregation Operators

There exist a great variety of aggregation operators. They can be roughly classified into several categories (Grabisch *et al.*, 1998), each possessing distinct behaviour or semantic: conjunctive, disjunctive, averaging (or compensative), non-compensative, and weighted operators, according to the way the values are aggregated.

Conjunctive operators are used to aggregate values when these values are orthogonal. The term “conjunctive” accents that, in this case, aggregation is in some way analogous to the logical operator “and” because the resulting value is high if and only if all the aggregated values are high. For binary operator this requirement can be expressed in the following axiom (Grabisch *et al.*, 1998):

$$\text{Aggreg}(1, a) = a, \quad \forall a \in [0, 1]. \quad (3.1)$$

It means that for conjunctive operators 1 is a neutral element.

Important sub-category of conjunctive operators is so called triangular norms (*t*-norms). Triangular norms are often denoted by *T*. They have been introduced by Schweizer and Sklar (Schweizer and Sklar, 1960; 1983) to model distance in probabilistic spaces. *t*-norm is a symmetric, associative, non-decreasing for any argument mapping

$$T: [0, 1]^n \rightarrow [0, 1], \quad (3.2)$$

for which 1 is neutral element. In aggregation theory *t*-norms are used to generalise Boolean logical operator “and” to multi-valued logic. For *t*-norms it is true that

$$T(x, y) \leq \min(x, y). \quad (3.3)$$

A *t*-norm is called strict if it is strictly increasing for any argument. Well-known examples of *t*-norms are operator $\min(x, y)$ and product xy . Axiomatic of *t*-norms attempts to capture the basic properties of set intersection. It is one of the main advantages of *t*-norms. The main disadvantages are that *t*-norms generally do not satisfy criteria (idempotence, compensativeness, scale invariance, etc.) required for the aggregation of values that have different nature, for example, for the aggregation results of measurements obtained using different metrics.

Disjunctive operators are dual of conjunctive operators. They are used in cases when aggregation operator must have properties analogous to the logical operator “or” or, in other words, when the resulting value must be high if and only if at least one of aggregated values is high. It will be low if and only if all aggregated values are low. For binary operator this requirement can be expressed in the following axiom (Grabisch *et al.*, 1998):

$$\text{Aggreg}(0, a) = a, \quad \forall a \in [0, 1]. \quad (3.4)$$

It means that for disjunctive operators 0 is a neutral element.

Important sub-category of disjunctive operators is so called triangular conorms (t -conorms, s -norms). Triangular conorms are often denoted by \perp . t -conorm is a symmetric, associative, non-decreasing for any argument mapping

$$\perp: [0, 1] \times [0, 1] \rightarrow [0, 1], \quad (3.5)$$

that satisfies the axiom (3.4). In aggregation theory t -conorms are used to generalise Boolean logical operator “or” to multi-valued logic. For t -conorms it is true that

$$\perp(x, y) \geq \max(x, y). \quad (3.6)$$

Using the construction

$$\perp(x, y) := 1 - T(1 - x, 1 - y) \quad (3.7)$$

for any t -norm T dual t -conorm \perp can be defined (Fullér, 2005). t -norm and a t -conorm are dual if they satisfy the DeMorgan law (Detyniecki, 2000):

$$\neg T(x, y) = \perp(\neg x, \neg y), \quad (3.8)$$

where negation usually is defined as a strong negation

$$\neg x = 1 - x. \quad (3.9)$$

Well-known examples of t -conorms are operator $\max(x, y)$ and probabilistic addition $x + y - xy$. Axiomatic of t -conorms attempts to capture the basic properties of set union. It is one of the main advantages of t -conorms. The main disadvantage is that t -conorms, similarly as t -norms cannot be used to aggregate the values of different nature.

For details of basic analytical and algebraic properties of t -norms and t -conorms see (Klement *et al.*, 2002).

The third category of aggregation operators is compensative operators that are neither conjunctive nor disjunctive. They are compensative in the sense that low values are compensated by high values and the result of combination is a medium value. Thus, compensative operators are averaging operators. They are monotonic, idempotent and are suitable for combining the values of different nature. Examples of compensative operators are mean operators¹, median and order statistics². A family of mean operators formed by different extensions of arithmetic mean is called quasi-arithmetic means. It is defined (Kolmogorov, 1930) as follows:

$$\text{Aggreg } f(x_1, x_2, \dots, x_n) = f^{-1} \left[\sum_{i=1}^n \frac{1}{n} f(x_i) \right], \quad (3.10)$$

¹Quasi-arithmetic means are idempotent, continuous, strictly monotonic, compensative and decomposable. They may have a neutral element. Arithmetic mean has the following additional property: it is stable with respect to linear changes of measurement scale.

²Median and statistics are idempotent, continuous, monotonic, compensative and stable with respect to linear changes of measurement scale. They may have a neutral element. In addition, medians are associative.

where f is any continuous strictly monotonic function.

Conjunctive, disjunctive and compensative operators are the main categories of aggregation operators. However, there exist some aggregation operators, so called non-compensative operators, which do not belong to any of these categories. Usually, non-compensative operators are more or less of average type. However, they may extend beyond minimum and maximum operators. Examples of non-compensative operators are symmetric sum (Silvert, 1979) and compensatory operators (Zimmermann and Zysno, 1980). Symmetric sums are continuous, monotonic and commutative. Compensatory operators are a mix of t -norm and t -conorm. Such mixed operators provide a kind of compensation of each other by the values that are aggregated. They have rather limited properties: continuity, monotonicity, and associativity. Some non-compensative operators have a neutral element.

All discussed categories of aggregation operators are non-weighted or, in other words, treat all arguments as of the same importance. However, in many cases arguments are not of the same importance. So, in such cases, the weights of arguments must be introduced. Generally, weights in aggregation can be necessary either from qualitative or quantitative reasons. Qualitative reasons arise when the values that are aggregated have different importance. Quantitative reasons arise in cases when the input of aggregation operator has different frequencies or cardinalities. Introduction of weights extends non-weighted operators. For example, minimum and maximum operators have been extended by Dubois and Prade (Dubois, 1985) to weighted minimum and weighted maximum operators:

$$w \min_{w_1, \dots, w_n} (x_1, x_2, \dots, x_n) = \min_{i=1}^n [\max(1 - w_i, x_i)], \quad (3.11)$$

$$w \max_{w_1, \dots, w_n} (x_1, x_2, \dots, x_n) = \max_{i=1}^n [\min(w_i, x_i)], \quad (3.12)$$

where weights $w_i \geq 0$ for all $i = 1, 2, \dots, n$ and $\max_{i=1}^n (w_i) = 1$.

These operators can be generalised to weighted t -norm T_w and weighted t -conorm \perp_w (Moraga, 2001). Let $w = [w_1, w_2, \dots, w_n]$, where $w_i \in [0, 1]$, $1 \leq i \leq n$. Then $\forall x_i \in [0, 1]$, $1 \leq i \leq n$

$$T_w(x_1, \dots, x_n) = T(\perp(x_1, 1 - w_1), \perp(x_2, 1 - w_2), \dots, \perp(x_n, 1 - w_n)), \quad (3.13)$$

$$\perp_w(x_1, \dots, x_n) = \perp(T(x_1, w_1), T(x_2, w_2), \dots, T(x_n, w_n)). \quad (3.14)$$

However, axiomatic definition of t -norms does not allow weighted aggregation. In order to obtain a weighted extension of t -norms, some of the axiomatic requirements must be dropped. Weighted operators by definition are not commutative, so commutativity and associativity axioms are eliminated for them.

EXAMPLE 1. Consider $T = xy$ is the probabilistic multiplication and $\perp = x + y - xy$ is the probabilistic addition. Then for $n = 2$ the weighted t -norm and the weighted t -conorm will be the following:

$$T_w(x_i, y_i) = (1 - w_1(1 - x_1))(1 - w_2(1 - x_2)), \quad (3.15)$$

$$\perp_w(x_i, y_i) = w_1x_1 + w_2x_2 - w_1w_2x_1x_2. \quad (3.16)$$

The whole family of quasi-arithmetic means (3.10) is extended by

$$\text{Aggreg}_f(x_1, x_2, \dots, x_n) = f^{-1} \left[\sum_{i=1}^n w_i f(x_i) \right], \quad (3.17)$$

where w_i are the weights of the values x_i .

EXAMPLE 2. Consider the arithmetic mean is extended to the weighted mean in the following way:

$$\text{Aggreg}(x_1, x_2, \dots, x_n) = \sum_{i=1}^n w_i x_i, \quad (3.18)$$

where the weights are normalised so that $\sum_{i=1}^n w_i = 1$.

As a natural framework for the inclusion of the behavioural properties of aggregation operators the family of ordered weighted operators (OWA) has been proposed (Yager, 1988). OWA operators are idempotent, monotonic and commutative. They are important because they integrate both conjunctive and disjunctive behaviour. The OWA family is defined as follows:

$$\text{Aggreg}_{\text{OWA}}(x_1, x_2, \dots, x_n) = \sum_{i=1}^n w_i \text{order}(i, x_1, x_2, \dots, x_n), \quad (3.19)$$

where function $\text{order}: \{1, 2, \dots, n\} \times [0, 1]^n \rightarrow [0, 1]$ defines ordering of the values x_1, x_2, \dots, x_n in increasing order³. Thus, in OWA operator the weight is associated with a particular ordered position of aggregate but not with a particular value (Fullér, 1996). The main advantage of OWA operators is their versatility (Detyniecki, 2000; Grabisch *et al.*, 1998). In case, when $w_1 = 1$ and all other weights are equal to zero, OWA is equivalent to maximum operator; in case, when only $w_n = 1$ and all other weights are equal to zero it is equivalent to minimum operator. Arithmetic mean also is a particular case of OWA operator. It is obtained when all weights are equal to $1/n$.

Further generalisation of average operators is so-called fuzzy integrals. It is evident that discretisation of traditional integral (including Lebesgue integral) of a function is a particular case of averaging operator because it represents the average value of this function (Grabisch *et al.*, 1998; Detyniecki, 2000). Sugeno (Sugeno, 1974) and other authors extended the notion of Lebesgue integrals by introducing the fuzzy measure that is considered as the weight of importance of a set of the values that are aggregated. It is defined as a mapping

$$m: P(V) \rightarrow [0, 1], \quad (3.20)$$

³Using substitution $n - i + 1$ it is possible to order the values in decreasing order.

where $V = \{v_1, v_2, \dots, v_n\}$ is a set of the values that are aggregated and $P(V)$ denotes the power set of V . In addition, the function m must satisfy boundary conditions and monotonicity axiom:

$$m(\emptyset) = 0, \quad m(V) = 1; \quad (3.21)$$

$$\forall A, B \in P(V) \left((A \subset B) \Rightarrow (m(A) \leq m(B)) \right). \quad (3.22)$$

Lebesgue integrals with fuzzy measure are called fuzzy integrals. Discrete fuzzy integrals can be seen as particular cases of averaging operators, too. There are two kinds of fuzzy integrals: Sugeno integrals and Choquet integrals. The discrete Choquet integral of values x_1, x_2, \dots, x_n for the values $V = \{v_1, v_2, \dots, v_n\}$ that are aggregated with respect to a fuzzy measure m is defined as follows:

$$\begin{aligned} \text{Aggreg}_{\text{Choquet}}(x_1, x_2, \dots, x_n) &= \sum_{i=1}^n (\text{order}(i, x_1, x_2, \dots, x_n) \\ &\quad - \text{order}(i-1, x_1, x_2, \dots, x_n)) m(\{x_i, x_{i+1}, \dots, x_n\}), \end{aligned} \quad (3.23)$$

where it is supposed that $\text{order}(0, x_1, x_2, \dots, x_n) = 0$.

In the case when fuzzy measure $m_0 = m(\emptyset) = 0$, and $m_i = m(S) = i$ for any $S \subseteq \{x_1, x_2, \dots, x_n\}$ of the cardinality i , the Choquet integral corresponds to the OWA operator:

$$\text{Aggreg}_{\text{OWA}}(x_1, x_2, \dots, x_n) = \sum_{i=1}^n (m_i - m_{i-1}) \text{order}(i, x_1, x_2, \dots, x_n). \quad (3.24)$$

And vice versa, each OWA operator corresponds to the Choquet integral with the measure $m_k = \sum_{i=k}^n w_i$.

The discrete Sugeno integral of values x_1, x_2, \dots, x_n for the values $V = \{v_1, v_2, \dots, v_n\}$ that are aggregated with respect to a fuzzy measure m is defined as follows:

$$\text{Aggreg}_{\text{Sugeno}} = \max_{i=1}^n \left(\min(\text{order}(i, x_1, x_2, \dots, x_n) m(x_i, x_{i+1}, \dots, x_n)) \right), \quad (3.25)$$

where it is supposed that $\text{order}(0, x_1, x_2, \dots, x_n) = 0$.

Sugeno integrals generalise the weighted minimum and the weighted maximum. In the case when fuzzy measure $m_0 = m(\emptyset) = 0$, and $m_i = m(S) = i$ for any $S \subseteq \{x_1, x_2, \dots, x_n\}$ of the cardinality i , the Sugeno integral corresponds to the weighted maximum:

$$\text{Aggreg}_{\text{max}} = \max_{i=1}^n \left(\min(\text{order}(i, x_1, x_2, \dots, x_n) m_i) \right). \quad (3.26)$$

In the case when fuzzy measure $m_0 = m(\emptyset) = 0$, and $m_i = m(S) = i$ for any $S \subseteq \{x_1, x_2, \dots, x_n\}$ of the cardinality i , the Sugeno integral corresponds to the weighted minimum:

$$\text{Aggreg}_{\min} = \min_{i=1}^n \left(\max(\text{order}(i, x_1, x_2, \dots, x_n) m_i) \right). \quad (3.27)$$

Choquet integrals and Sugeno integrals have been widely studied in literature and compared to each other (see, for example (Bolanos *et al.*, 1996)). These integrals are complementary in the sense that Choquet integrals can be seen as a kind of average, while Sugeno integrals can be seen as a kind of median. Thus, each integral is intended to be used for different purposes.

4. Aggregation of Quality Characteristics

Let us consider further how to apply surveyed above aggregation theory in order to solve the problem of aggregation of quality characteristics of an IS specification language.

It is highly desirable that internal quality of a language will be measured using some numerical scale. It is also preferable to use a scale of cardinal type. On the other hand, internal quality of a language cannot be measured in any standard units, like meters or amperes, of some measurement system. In addition, internal quality characterises the potential applicability of a language only but says nothing about its quality in use or about the degree to which this language satisfies the requirements of a particular project. These are the main reasons to define internal quality of a language through the probability that this language will satisfy the requirements of any possible project regardless of its complexity or size (Caplinskas and Gasperovic, 2005b). If, for example, quality of some specification language is characterised by the value 0.8, it means that this language will be acceptable approximately for 4/5 of all imaginable projects.

The quality is described in a hierarchical way: it is decomposed into characteristics and sub-characteristics of several levels. Each characteristic (sub-characteristic) describes some feature of a language. Characteristics of the lowest level are described by probabilities that corresponding feature of a language will be sufficient for any possible project. So, the notion of probability allows expressing partiality that is natural for specification languages, because almost always a language supports any particular feature at some limited extent only. Probabilities can be evaluated, for example, using some assessment case suites, however, this paper does not address the problem how to evaluate probabilities describing the completeness of the bottom-level features of a specification language.

The values of the characteristics of all higher levels can be calculated using some aggregation techniques. Let us consider this problem in detail.

Let us denote the feature of the language L described by characteristic ξ by $L(\xi)$, the probability that $L(\xi)$ will be sufficient for any theoretically imaginable project by $p(\xi)$, the probability that $L(\xi)$ will become necessary for some project P by $q(\xi)$. Let us discuss now the techniques to aggregate measurements of sub-characteristics.

There are several cases and for each of them different aggregation method should be used:

1. All sub-characteristics $\xi_1, \xi_2, \dots, \xi_n$ of the characteristic ξ are orthogonal and for $\forall i \in [1, n]$ ($q(\xi_i) = 1$).
2. All sub-characteristics $\xi_1, \xi_2, \dots, \xi_n$ of the characteristic ξ are orthogonal and there exists such subset $\xi' \subseteq \{\xi_1, \xi_2, \dots, \xi_n\}$ that $\xi' = \{\xi_k | q(\xi_k) < 1\}$.
3. There exists some main sub-characteristic of the characteristic ξ , say ξ_1 , in the sense that $q(\xi_1) = 1$. For all other sub-characteristics ξ_2, \dots, ξ_n is true that $q(\xi_i) < 1$ for $i \in [2, n]$ and $L(\xi_2) \text{ suppl } L(\xi_1)$, $L(\xi_3) \text{ suppl } L(\xi_2)$, \dots , $L(\xi_n) \text{ suppl } L(\xi_{n-1})$, where $L(\xi_k) \text{ suppl } L(\xi_l)$ means that $L(\xi_k)$ is supplementary (additional) for $L(\xi_l)$.
4. All sub-characteristics $\xi_1, \xi_2, \dots, \xi_n$ of the characteristic ξ are alternative, $q(\xi_i) < 1$ for any $i \in [1, n]$ and is true that:

$$\sum_{i=1}^n q(\xi_i) = 1. \quad (4.1)$$

Because the measures of sub-characteristics are probabilities $p(\xi_1), p(\xi_2), \dots, p(\xi_n)$, the measure of the characteristic ξ can be regarded as compound probability and methods of the probability theory can be used for aggregation. It means that probability that $L(\xi)$ will be sufficient for any possible project P depends on the probability that any $L(\xi_i)$, where $i \in [1, n]$, will be sufficient for this aim. Probability theory provides two ways to combine probabilities (Pfeiffer, 1979): multiplication rule and addition rule. We will apply the multiplication rule to aggregate orthogonal sub-characteristics or, in other words, in cases when features described by these sub-characteristics are used for different aims. Addition rule will be applied to aggregate supplemental and alternative sub-characteristics or, in other words, when features described by these sub-characteristics are used for the same aim and compensate each other or are alternative to each other. Let us consider now how to apply multiplication and addition rules for aggregation of quality sub-characteristics in each of cases mentioned above. We will also check for each case that the proposed aggregation technique is an aggregation operator and that it indeed has the properties required in this particular case to aggregate sub-characteristics in a proper way.

In the first case all sub-characteristics are orthogonal. If two sub-characteristics ξ_1 and ξ_2 are orthogonal it means that $L(\xi_1)$ and $L(\xi_2)$ are independent, used for different aims. So, in the first case all characteristics are independent and will become necessary (with the probability 1) for any possible project. Consequently, in this case the multiplication rule can be applied directly to aggregate the sub-characteristics.

$$p(\xi) = \prod_{i=1}^n p(\xi_i). \quad (4.2)$$

In order to illustrate the proposed aggregation techniques, we use geometric probability. As sample space (i.e., the set of all possible values the events may assume) we use

unitary square. Event space (i.e., the subset of the sample space consisting of events that represent a successful outcome) is represented by the corresponding area in the unitary square. Further we use Venn diagrams to illustrate aggregation techniques for aggregation of some small number of sub-characteristics. Fig. 1 represents the result of aggregation of $p(\xi_1)$ and $p(\xi_2)$ by the area $p(\xi_1)p(\xi_2)$.

Statement 1. *The formula (4.2) defines a t-norm.*

Proof. It is evident that operator $\prod_{i=1}^n p(\xi_i)$ is aggregation operator. Satisfaction of axioms (2.1)–(2.3) follows immediately from the properties of probabilistic multiplication operator:

- for $n = 1$, the formula (4.2) defines identity (i.e., $p(\xi) = p(\xi_1)$);
- boundary conditions are satisfied: because (if $p(\xi_i) = 0$ for any i then $p(\xi) = 0$); if $p(\xi_i) = 1$ for any i then $p(\xi) = 1$);
- multiplication is monotonic in the interval $[0,1]$.

It is also evident that probabilistic multiplication is symmetric, associative and that for probabilistic multiplication 1 is the neutral element. Consequently, (4.2) defines a t-norm.

Thus, the aggregation technique defined by formula (4.2) really can be applied to aggregate orthogonal sub-characteristics in a proper way, since t-norms are conjunctive aggregation operators generalising Boolean logical operator “and” to multi-valued logic.

EXAMPLE 3. Internal quality of a specification language L is defined through four characteristics: functionality ξ_1 , reliability ξ_2 , usability ξ_3 , and efficiency ξ_4 . Measures of these characteristics are probabilities $p(\xi_1)$, $p(\xi_2)$, $p(\xi_3)$ and $p(\xi_4)$ correspondingly. All four characteristics are orthogonal to each other (i.e., describes the features of a language that are used for different purposes) and the probabilities $q(\xi_1)$, $q(\xi_2)$, $q(\xi_3)$ and $q(\xi_4)$ that they will become necessary for any project P are equal to 1. Thus internal quality of the language L is described by the probability $p(\xi) = p(\xi_1)p(\xi_2)p(\xi_3)p(\xi_4)$.

In the second case, all sub-characteristics $\xi_1, \xi_2, \dots, \xi_n$ of the characteristic ξ are also orthogonal but there exists at least one sub-characteristic ξ_i for which $q(\xi_i) < 1$. In this case multiplication rule cannot be applied directly. It must be changed in the way to meet the requirement that any feature $L(\xi)$ must be sufficient for any particular project at the

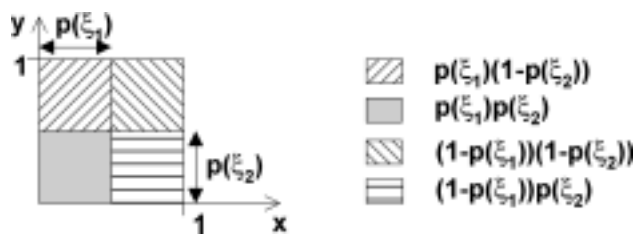


Fig. 1. Aggregation of orthogonal quality characteristics required for any project (case 1).

same degree as its sub-features $L(\xi_i)$ are sufficient for this project in case when these sub-features are required for this project at all. Of course, $L(\xi_i)$ is sufficient also for projects for which it is not required. It means that multiplication rule must be modified to cover also the projects for which $L(\xi)$ is not required:

$$p(\xi) = \prod_{i=1}^n \left(1 - q(\xi_i)(1 - p(\xi_i))\right). \quad (4.3)$$

Fig. 2 represents the result of aggregation of one sub-characteristic, which is evaluated as $p(\xi_1)$ and for which $q(\xi_1) < 1$. In this case aggregated characteristic is expressed by the area $p(\xi_1)q(\xi_1)$ (for the case that characteristic is indeed sufficient for any possible project) and areas $p(\xi_1)(1 - q(\xi_1))$ and $(1 - p(\xi_1))(1 - q(\xi_1))$ (for the case, when the characteristic is not required for this project at all). Although only one characteristic is illustrated in Fig. 2, this principle can be generalised for any number of quality sub-characteristics.

Statement 2. *The formula (4.3) defines a weighted t-norm.*

Proof. Rewrite formula (4.3) in the form

$$p(\xi) = (1 - q(\xi_1)(1 - p(\xi_1)))(1 - q(\xi_2)(1 - p(\xi_2))) \dots (1 - q(\xi_n)(1 - p(\xi_n))).$$

Denote $T = xy$ and $\perp = x + y - xy$. Let us use now proof by induction.

For $n = 2$

$$\begin{aligned} p(\xi) &= \left(1 - q(\xi_1)(1 - p(\xi_1))\right) \left(1 - q(\xi_2)(1 - p(\xi_2))\right) \\ &= (1 - q(\xi_1) + q(\xi_1)p(\xi_1))(1 - q(\xi_2) + q(\xi_2)p(\xi_2)) \\ &= (1 + p(\xi_1) - p(\xi_1) - q(\xi_1) + q(\xi_1)p(\xi_1))(1 + p(\xi_2) \\ &\quad - p(\xi_2) - q(\xi_2) + q(\xi_2)p(\xi_2)) \\ &= (p(\xi_1) + 1 - q(\xi_1) - p(\xi_1)(1 - q(\xi_1)))(p(\xi_2) + 1 - q(\xi_2) - p(\xi_2)(1 - q(\xi_2))) \\ &= \perp(p(\xi_1), 1 - q(\xi_1)) \perp(p(\xi_2), 1 - q(\xi_2)) \\ &= T(\perp(p(\xi_1), 1 - q(\xi_1)), \perp(p(\xi_2), 1 - q(\xi_2))). \end{aligned}$$



Fig. 2. Aggregation of orthogonal quality characteristics that are not all required for any project (case 2).

Thus, for $n = 2$ the statement 2 is true.

Now suppose that the statement 2 is true for $n = i$. It means that for $n = i$

$$p(\xi) = (1 - q(\xi_1))(1 - p(\xi_1))(1 - q(\xi_2))(1 - p(\xi_2)) \dots (1 - q(\xi_i))(1 - p(\xi_i))$$

defines weighted t -norm

$$T\left(\perp(p(\xi_1), 1 - q(\xi_1)), \perp(p(\xi_2), 1 - q(\xi_2)), \dots, \perp(p(\xi_i), 1 - q(\xi_i))\right).$$

For $i = i + 1$

$$\begin{aligned} p(\xi) &= T\left(\perp(p(\xi_1), 1 - q(\xi_1)), \perp(p(\xi_2), 1 - q(\xi_2)), \dots, \right. \\ &\quad \left. \perp(p(\xi_i), 1 - q(\xi_i))\right)(1 - q(\xi_{i+1})(1 - p(\xi_{i+1})) \\ &= (1 - q(\xi_1))(1 - p(\xi_1))(1 - q(\xi_2))(1 - p(\xi_2)) \dots \\ &\quad (1 - q(\xi_i))(1 - p(\xi_i))(1 - q(\xi_{i+1})(1 - p(\xi_{i+1})) \\ &= T\left(\perp(p(\xi_1), 1 - q(\xi_1)), \perp(p(\xi_2), 1 - q(\xi_2)), \dots, \right. \\ &\quad \left. \perp(p(\xi_i), 1 - q(\xi_i)), \perp(p(\xi_{i+1}), 1 - q(\xi_{i+1}))\right). \end{aligned}$$

Consequently, (4.3) defines weighted t -norm $T(\perp(p(\xi_1), 1 - q(\xi_1)), \perp(p(\xi_2), 1 - q(\xi_2)), \dots, \perp(p(\xi_n), 1 - q(\xi_n)))$.

Thus, the aggregation technique defined by formula (4.3) really can be applied to aggregate orthogonal sub-characteristics of different importance in a proper way, since weighted t -norms are theoretically sound extension of t -norms.

EXAMPLE 4. Semantic sufficiency ξ of a specification language L (Caplinskas and Gasperovic, 2005b) characterises the conceptual level of the linguistic system Φ_L beyond this language. It is the measure of the ability of language L to specify all “things” that might be necessary for analysis and design of any possible project P . Semantic sufficiency ξ is decomposed into ontological sufficiency ξ_1 and epistemological sufficiency ξ_2 . Ontological sufficiency ξ_1 is characterised through the probability $p(\xi_1)$ that any project P will be conceptualised successfully through categories α_L provided by the linguistic system Φ_L . Epistemological sufficiency ξ_2 characterises the ability of the linguistic system Φ_L to express epistemological primitives or, in other words, characterises the constructive power of “algebra of concepts” provided by Φ_L . ξ_2 is characterised through the probability $p(\xi_2)$ that all required conceptual structures will be modelled using constructs of language L . The probability $q(\xi_1)$ that ξ_1 will become necessary for any project P is equal to 1 and the probability $q(\xi_2)$ that ξ_2 will become necessary for any project P is less than 1. Thus the measure of semantic sufficiency ξ is described by the probability $p(\xi) = (1 - q(\xi_1))(1 - p(\xi_1))(1 - q(\xi_2))(1 - p(\xi_2)) = p(\xi_1)(1 - q(\xi_2)(1 - p(\xi_1)))$.

In the third case sub-characteristics are supplemental to the main sub-characteristic ξ_1 or, in other words, all features $L(\xi_1), L(\xi_2), \dots, L(\xi_n)$ are used for the same purpose and supplemental features are necessary only when previous features are insufficient for this purpose. In this case addition rule can be applied, however, it must be modified appropriately because $q(\xi)$ can be expressed through $p(\xi)$ in the following way:

$$\begin{aligned} q(\xi_1) &= 1, \\ q(\xi_2) &= 1 - p(\xi_1), \\ q(\xi_3) &= 1 - p(\xi_1) - p(\xi_2)(1 - p(\xi_1)), \\ q(\xi_4) &= 1 - p(\xi_1) - p(\xi_2) - p(\xi_3)\left(1 - p(\xi_1) - p(\xi_2)(1 - p(\xi_1))\right), \\ &\dots \\ q(\xi_n) &= 1 - p(\xi_1) - \dots - p(\xi_{n-1})q(\xi_{n-1}). \end{aligned}$$

Thus,

$$\begin{aligned} p(\xi) &= \sum_{i=1}^n p(\xi_i)q(\xi_i) = p(\xi_1) + p(\xi_2)q(\xi_2) + p(\xi_3)q(\xi_3) + \dots + p(\xi_n)q(\xi_n) \\ &= \sum_{1 \leq i \leq n} p(\xi_i) - \sum_{1 \leq j < k \leq n} p(\xi_j)p(\xi_k) + \sum_{1 \leq j < k < s \leq n} p(\xi_j)p(\xi_k)p(\xi_s) - \dots \\ &\quad + (-1)^{n-1} p(\xi_1) \dots p(\xi_n) \\ &= \sum_{r=1}^n (-1)^{r-1} \sum_{1 \leq j < k < s < \dots < r \leq n} p(\xi_1) \dots p(\xi_r). \end{aligned} \quad (4.4)$$

This formula is well-known inclusion-exclusion formula (Durrett, 1993). Fig. 3 represents the result of aggregation of $p(\xi_1)$, $p(\xi_2)$ and $p(\xi_3)$. In this case from the area $p(\xi_1) + p(\xi_2) + p(\xi_3)$ it is necessary to exclude the areas $p(\xi_1)p(\xi_3)$, $p(\xi_2)p(\xi_3)$, $p(\xi_1)p(\xi_2)$ and to include the area $p(\xi_1)p(\xi_2)p(\xi_3)$, because when $n > 2$ the exclusion of the pairwise intersections is too severe, thus compensating inclusion is required.

Statement 3. The formula (4.4) defines *t*-conorm.

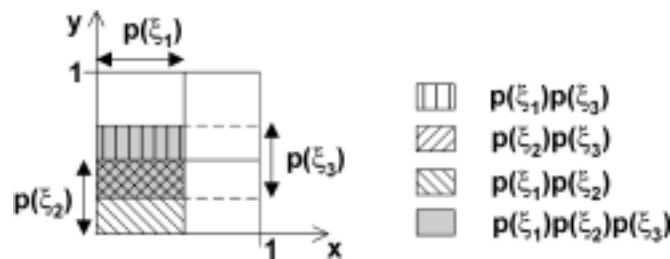


Fig. 3. Aggregation of supplemental quality characteristics (case 3).

Proof. Denote $\perp = x + y - xy$. Let us use now proof by induction.

For $n = 2$

$$p(\xi) = p(\xi_1) + p(\xi_2) - p(\xi_1)p(\xi_2) = \perp(p(\xi_1), p(\xi_2)).$$

Thus, for $n = 2$ the statement 3 is true.

Now suppose that the statement 3 is true for $n = m$. It means that for $n = m$

$$p(\xi) = p(\xi_1) + p(\xi_2) - p(\xi_1)p(\xi_2) - \dots + (-1)^{m-1}p(\xi_1) \dots p(\xi_m)$$

defines t -conorm

$$\perp(p(\xi_1), p(\xi_2), \dots, p(\xi_m)).$$

For $m = m + 1$

$$\begin{aligned} p(\xi) &= \perp(p(\xi_1), p(\xi_2), \dots, p(\xi_m))(-1)^m p(\xi_1) \dots p(\xi_{m+1}) \\ &= \sum_{1 \leq i \leq m} p(\xi_i) - \sum_{1 \leq j < k \leq m} p(\xi_j)p(\xi_k) + \sum_{1 \leq j < k < s \leq m} p(\xi_j)p(\xi_k)p(\xi_s) - \dots \\ &\quad + (-1)^{m-1} p(\xi_1) \dots p(\xi_m) + p(\xi_m) - \sum_{1 \leq i \leq m} p(\xi_i)p(\xi_{m+1}) \\ &\quad + \sum_{1 \leq j < k \leq m} p(\xi_j)p(\xi_k)p(\xi_{m+1}) - \sum_{1 \leq j < k < s \leq m} p(\xi_j)p(\xi_k)p(\xi_s)p(\xi_{m+1}) \\ &\quad + \dots + (-1)^{m+1} p(\xi_1) \dots p(\xi_m)p(\xi_{m+1}) \\ &= \sum_{1 \leq i \leq m+1} p(\xi_i) - \sum_{1 \leq j < k \leq m+1} p(\xi_j)p(\xi_k) \\ &\quad + \sum_{1 \leq j < k < s \leq m+1} p(\xi_j)p(\xi_k)p(\xi_s) - \dots + (-1)^m p(\xi_1) \dots p(\xi_{m+1}) \\ &= \perp(p(\xi_1), p(\xi_2), \dots, p(\xi_{m+1})). \end{aligned}$$

Consequently, (4.4) defines t -conorm $\perp(p(\xi_1), p(\xi_2), \dots, p(\xi_n))$.

Thus, the aggregation technique defined by formula (4.4) really can be applied to aggregate in a proper way such non-orthogonal sub-characteristics, for which it is true that at least one sub-characteristic will become necessary for any possible project with the probability 1, since t -conorms generalise Boolean logical operator “or” to multi-valued logic.

EXAMPLE 5. Functionality ξ of a specification language L (Caplinskas and Gasperovic, 2005b) is defined as the existence of means required specifying functional and non-functional properties of the subject system. Functionality has two sub-characteristics:

suitability ξ_1 and flexibility ξ_2 . The sub-characteristic ξ_1 is main sub-characteristic. It characterises how sophisticated statements about potential systems in a particular realm a specification language L is able to express, and at what level of granularity it can be done. The measure of ξ_1 is the probability $p(\xi_1)$ that any statement about any system of a particular kind can be formulated in terms of the language L . The probability becomes equal to 1, if the language L allows formulating statements about any property of a system in a given realm and expressing it with the needed degree of precision. The sub-characteristic ξ_2 is supplemental. It describes the extent to which the language can be adjusted to specify preliminary not intended properties. The measure of ξ_2 is the probability $p(\xi_2)$ that the language L can be applied to the whole spectrum of IS-related systems, namely, business systems (from the IS design perspective), information systems, and applications. The probability $q(\xi_2) = 1 - p(\xi_1)$, because ξ_2 becomes necessary only in cases when $L(\xi_1)$ is not sufficient to specify the subject system. Thus, $p(\xi) = p(\xi_1) + p(\xi_2)q(\xi_2) = p(\xi_1) + p(\xi_2)(1 - p(\xi_1)) = p(\xi_1) + p(\xi_2) - p(\xi_1)p(\xi_2)$.

In the fourth case all sub-characteristics $\xi_1, \xi_2, \dots, \xi_n$, are alternative. It means that there exist no main sub-characteristic or, in other words, $q(\xi_i) < 1$ for any $i \in [1, n]$, however, at least one of features $L(\xi_i)$ must be used. Thus the condition (4.1) must be satisfied. In this case addition rule must be changed taking into account that sub-characteristics may duplicate each other and that the overlapping of alternative sub-characteristics should be eliminated:

$$p(\xi) = \sum_{i=1}^n p(\xi_i)q(\xi_i) - \prod_{i=1}^n p(\xi_i)q(\xi_i). \tag{4.5}$$

Fig. 4 represents the result of aggregation of $p(\xi_1)$ and $p(\xi_2)$. In this case from the area $p(\xi_1)q(\xi_1) + p(\xi_2)q(\xi_2)$ it is necessary to exclude the area $p(\xi_1)q(\xi_1)p(\xi_2)q(\xi_2)$.

Statement 4. *The formula (4.5) defines a weighted t-conorm.*

Proof. Rewrite formula (4.5) in the form

$$p(\xi) = q(\xi_1)p(\xi_1) + q(\xi_2)p(\xi_2) + \dots + q(\xi_n)p(\xi_n) - q(\xi_1)p(\xi_1)q(\xi_2)p(\xi_2) \dots q(\xi_n)p(\xi_n).$$

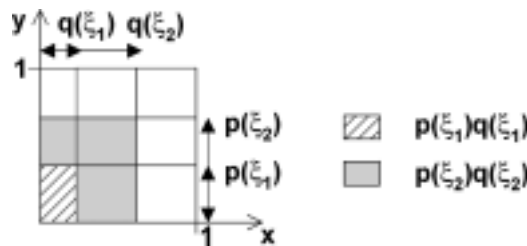


Fig. 4. Aggregation of alternative quality characteristics (case 4).

Denote $T = xy$ and $\perp = x + y - xy$. Let us use now proof by induction.

For $n = 2$

$$\begin{aligned} p(\xi) &= q(\xi_1)p(\xi_1) + q(\xi_2)p(\xi_2) - q(\xi_1)p(\xi_1)q(\xi_2)p(\xi_2) \\ &= \perp(q(\xi_1)p(\xi_1), q(\xi_2)p(\xi_2)) = \perp\left(T(p(\xi_1), q(\xi_1)), T(p(\xi_2), q(\xi_2))\right). \end{aligned}$$

Thus, for $n = 2$ the statement 4 is true.

Now suppose that the statement 4 is true for $n = i$. It means that for $n = i$

$$\begin{aligned} p(\xi) &= q(\xi_1)p(\xi_1) + q(\xi_2)p(\xi_2) + \dots + q(\xi_i)p(\xi_i) \\ &\quad - q(\xi_1)p(\xi_1)q(\xi_2)p(\xi_2) \dots q(\xi_i)p(\xi_i) \end{aligned}$$

defines weighted t -conorm

$$\perp\left(T(p(\xi_1), q(\xi_1)), T(p(\xi_2), q(\xi_2)), \dots, T(p(\xi_i), q(\xi_i))\right).$$

For $i = i + 1$

$$\begin{aligned} p(\xi) &= \perp\left(T(p(\xi_1), q(\xi_1)), T(p(\xi_2), q(\xi_2)), \dots, T(p(\xi_i), q(\xi_i))\right)q(\xi_1)p(\xi_1) \\ &\quad + q(\xi_2)p(\xi_2) + \dots + q(\xi_i)p(\xi_i) + q(\xi_{i+1})p(\xi_{i+1}) \\ &\quad - q(\xi_1)p(\xi_1)q(\xi_2)p(\xi_2) \dots q(\xi_i)p(\xi_i)q(\xi_{i+1})p(\xi_{i+1}) \\ &= \perp\left(T(p(\xi_1), q(\xi_1)), T(p(\xi_2), q(\xi_2)), \dots, T(p(\xi_i), q(\xi_i)), T(p(\xi_{i+1}), q(\xi_{i+1}))\right). \end{aligned}$$

Consequently, (4.5) defines weighted t -conorm $\perp(T(p(\xi_1), q(\xi_1)), T(p(\xi_2), q(\xi_2)), \dots, T(p(\xi_n), q(\xi_n)))$.

Thus, the aggregation technique defined by formula (4.5) really can be applied to aggregate in a proper way non-orthogonal sub-characteristics of different importance, since weighted t -conorms are theoretically sound extension of t -conorms.

EXAMPLE 6. Flexibility ξ of the language L (Caplinskas and Gasperovic, 2005b) has three sub-characteristics: universality ξ_1 , adaptability ξ_2 , and extensibility ξ_3 . These characteristics are alternative. Sub-characteristic ξ_1 characterises the degree of generality of ontological primitives beyond the language L . The measure of ξ_1 is the probability $p(\xi_1)$ that ontological primitives, provided by categories α_L of the linguistic system Φ_L , are suitable to model concepts from the whole spectrum of subject systems. Sub-characteristic ξ_2 describes the ability of the language L to configure syntax and semantics to adapt it for arbitrary domain. The measure of ξ_2 is the probability $p(\xi_2)$ that any required specific of domain can be introduced using adaptability mechanisms provided by the language L . ξ_3 is closely related to ξ_2 . Although the features $L(\xi_2)$ and $L(\xi_1)$ in some sense are opposite to each other, we can consider them as alternative, because it is possible to implement flexibility through universality

as well as through adaptability. Sub-characteristic ξ_3 is the characteristics of the language that bears on its ability to define new constructs. The measure of extensibility ξ_3 is the probability $p(\xi_3)$ that any required domain-specific features can be introduced using extensibility mechanisms provided by the language L . Both ξ_2 and ξ_3 are important for general-purpose languages only. If the probability that ξ_1 will become necessary for any project P is $q(\xi_1)$, the probability that ξ_2 will become necessary for any project P is $q(\xi_2)$ and the probability that ξ_3 will become necessary for any project P is $q(\xi_3)$, then flexibility ξ of the language L can be described by the probability $p(\xi) = p(\xi_1)q(\xi_1) + p(\xi_2)q(\xi_2) + p(\xi_3)q(\xi_3) - p(\xi_1)q(\xi_1)p(\xi_2)q(\xi_2)p(\xi_3)q(\xi_3)$.

5. Aggregation of Measurements

Usually, it is very difficult or even impossible to develop precise metrics to measure the characteristics of internal quality of a specification language. One way to measure quality characteristics is to use representative case suits. However, the notion of representativity cannot be defined strictly and precisely. It is defined on an empirical basis. So, in order to be more precise, any sub-characteristic ξ_i of the internal quality of a language L should be measured using a number of different metrics (case suits) and results of different measurements should be aggregated. As a rule, three different case suits are used for this aim. Ideally, all measurements should produce almost identical results because all test cases should be debugged using a number of most popular specification languages (Z, UML, etc.). However, when a new specification language is evaluated it may occur that it accents a feature that is secondary only in most popular languages and for this reason has been underestimated in some test case suits. Thus, in some cases an unacceptable dispersion of measurement values may appear. Unacceptable dispersion can also be generated by inaccuracy of measurements, because test cases are specified manually and, consequently, measurement results can be impacted by the errors of the particular specifier. There is no way to determine the source of dispersion. It means that in both situations the same approach should be used to process unacceptable dispersion.

If the dispersion produces obviously meaningless results, for example, out of the interval $[0,1]$, the only way to solve this problem is to analyse all test cases, to discover reasons of such behaviour and to improve appropriate case suit or to perform more accurate measurements. However, it is very expensive solution and can be accepted in some exceptional cases only. In most of the cases, it is preferable to minimise the dispersion in some way or, in other words, to use appropriate weights for aggregation of measurement values. Partially, it is because the taxonomy of characteristics encompasses fifty seven characteristics and has five hierarchical levels and, therefore, the impact of one or even several characteristics to the final result of aggregation is very small.

Different approaches can be used to minimise possible deviations of measurement results. Most of methods of smoothing over measurement errors are based on the assumption that measurement errors are random errors and that measurements obey a normal, or Gaussian, distribution. It means that the measurements should be distributed about the

mean in such a manner that more measurements lie close to the mean than lie far away from it. We also accept this assumption. However, we cannot use so called Q-tests, empirical schemes that are used in statistics for testing inaccurate values (so called outliers) in small data sets, for example, Grubbs Test (EPA, 1992), Dixon Test (EPA, 1996), or Discordance Test (EPA, 1996), because in our case not only the deviations from Gaussian distribution, but also the distribution of values at some extent is unacceptable. In addition, the reliability of Q-tests for data sets including only three values is very low, although theoretically all mentioned above tests can be applied for such data sets, too. Aggregation operators with static weights (for example, Choquet integral) also cannot be used, because they require analysis of each particular case to assign appropriate weights. So, we propose the following heuristic, in which weights for aggregation of measurement values are determined dynamically:

- 1) to calculate the arithmetic mean

$$M(x_1, x_2, \dots, x_n) = \frac{1}{n} \sum_{i=1}^n x_i = \sum_{i=1}^n \left(\frac{1}{n} \cdot x_i \right) \quad (5.1)$$

and the standard deviation

$$d = \sqrt{\frac{\sum_{i=1}^n (x_i - M(x_1, x_2, \dots, x_n))^2}{n}} \quad (5.2)$$

for the set of measurements $\{x_1, x_2, \dots, x_n\}$ of the characteristic ξ ;

- 2) to substitute the values that are out of interval $[M - d; M + d]$ by the value $M - d$ or $M + d$ correspondingly;
- 3) to calculate the new arithmetic mean M_1 and to take it as the value of characteristic ξ .

This heuristic can be seen as a kind of combination of arithmetic and winsorised means. It allows dynamic determining of weights and minimising of dispersion. The main advantage of the proposed heuristic is that it can be used in all cases when the dispersion of the values is not obviously meaningless, including cases when two measurements are equally distant from the mean (in this case the heuristic rejects both the highest and the lowest values) and even cases when the dispersion is at acceptable extent.

EXAMPLE 7. Ontological sufficiency ξ of a language L was measured using metrics $\mu = \{m_1, m_2, m_3\}$. Measurements resulted in three different results: $m_1(\xi) = 0.1$; $m_2(\xi) = 0.2$; $m_3(\xi) = 0.9$. So, the aggregation of measurements should be done in the following way:

1. Firstly, the arithmetic mean of measurements

$$M = \frac{1}{3}(0.1 + 0.2 + 0.9) = 0.4$$

and the standard deviation

$$d = \sqrt{\frac{1}{3}((0.1 - 0.4)^2 + (0.2 - 0.4)^2 + (0.9 - 0.4)^2)} = 0.3559$$

are calculated by the formulas (5.1) and (5.2).

2. Secondly, the value 0.9 that is out of the interval

$$[M - d; M + d] = [0.0441; 0.7559]$$

is substituted by the value 0.7559.

3. Finally, the new arithmetic mean is calculated

$$M_1 = \frac{1}{2}(0.1 + 0.2 + 0.7559) = 0.35197$$

is calculated and the value 0.35197 is taken as the value of the characteristic ξ .

EXAMPLE 8. Let three different measurements of the epistemological sufficiency ξ of a language L using metric m have given results with small dispersion: $m_1(\xi) = 0.81$; $m_2(\xi) = 0.82$; $m_3(\xi) = 0.89$. Using the proposed heuristic to aggregate measurements, we obtain the following results:

1. The arithmetic mean

$$M = \frac{1}{3}(0.81 + 0.82 + 0.89) = 0.84$$

and the standard deviation

$$d = \sqrt{\frac{1}{3}((0.81 - 0.84)^2 + (0.82 - 0.84)^2 + (0.89 - 0.84)^2)} = 0.0356$$

are calculated by the formulas (5.1) and (5.2).

2. The value 0.89 that is out of the interval $[M - d; M + d] = [0.8044; 0.8756]$ is replaced by the value 0.8756.

3. New arithmetic mean

$$M_1 = \frac{1}{3}(0.81 + 0.82 + 0.8756) = 0.8352$$

is calculated and the value 0.8352 is taken as the value of characteristic ξ .

The presented examples demonstrate that the heuristic produces acceptable results in the case with large dispersion as well as in case with small dispersion.

6. Conclusions

Internal quality of a specification language L describes the quality that is independent from any context of use. Because of the imprecise nature of quality characteristics, it is reasonable to define such quality as the probability that the language L will satisfy the requirements of any imaginable project P . In an analogous way can be defined also all characteristics of internal quality, including elementary ones. Because characteristics of internal quality of the language L form a large hierarchical structure F , it is not obvious how to aggregate properly sub-characteristics through the whole structure F . Techniques of aggregation depend on the kind of dependences among characteristics that are aggregated. There exists four kinds of such dependencies: characteristics are orthogonal (independent) and all are required for any project; characteristics are orthogonal but not all are required for any project; characteristics supplement the one that is required for any project; and none of the characteristics is required for any project. In the first case the characteristics can be aggregated properly using a kind of t -norm, in the second case using a kind of weighted t -norm, in the third case using a kind of t -conorm, and in the fourth case using a kind of weighted t -conorm. In order to minimise possible deviations generated by shortcomings of the particular metric or by inaccuracy of the particular measurement, the mean with weights that are determined dynamically should be calculated. For this aim the paper proposes some heuristic, which can be seen as a kind of combination of arithmetic and winsorised means.

References

- Adomenas, P.G., and A. Ciucelis (2002). Data aggregation sets in adaptive data model. *Informatica*, **13**(4), 381–392.
- Bolanos, M.J., L.D. Hernandez and A. Salmeron (1996). Numerical experimentation and comparison of fuzzy integrals. *Mathware & Soft Computing*, **3**(5), 309–319.
- Caplinskas, A., and J. Gasperovic (2005a). An approach to evaluate quality in use of IS specification language. In J. Barzdins and A. Caplinskas (Eds.), *Frontiers in Artificial Intelligence and Applications*, vol. 118. IOS Press, Amsterdam. pp. 152–166.
- Caplinskas, A., and J. Gasperovic (2005b). Functionality of information systems specification language: concept, evaluation methodology, and evaluation problems. In O. Vasilecas, A. Caplinskas, W. Wojtkowski, W.G. Wojtkowsky, S. Wrycza and J. Zupancic (Eds.), *Information Systems Development. Advances in Theory, Practice and Education*. Kluwer Plenum Publishers. pp. 341–351.
- Caplinskas, A., A. Lupeikiene and O. Vasilecas (2002). A framework to analyse and evaluate Information Systems specification languages. *Lecture Notes in Computer Science*, **2435**, 248–262.
- Detyniecki, M. (2000). *Mathematical Aggregation Operators and their Application to Video Querying*. Phd. thesis, Universite Pierre & Marie Curie.
- Dubois, D., and H. Prade (1985). A review of fuzzy set aggregation connectives. *Information Sciences*, **36**, 85–121.
- Durrett, R. (1993). *The Essentials of Probability (Statistics)*. Duxbury Press.
- Environmental Protection Agency (1992). *Statistical Training Course for Ground-Water Monitoring Data Analysis, EPA/530-R-93-003*. Office of Solid Waste, Washington, DC.
- Environmental Protection Agency (1996). *Guidance for Data Quality Assessment, EPA/600-R-96-084*. Office of Research and Development, Washington, DC.
- Fullér, R. (1996). OWA operators in decision making. In C. Carlsson (Ed.), *Exploring the Limits of Support Systems*, TUCS General Publications, No. 3, Turku Center for Computer Science, Åbo. pp. 85–104.

- Fuller, R. (2005). *Neural Fuzzy Systems*. Lectures. Abo Academy.
<http://www.abo.fi/~rfuller/nfs.html>
- Grabisch, M., S.A. Orlovski and R.R. Yager (1998). Fuzzy aggregation of numerical preferences. In R. Slowinski (Ed.), *The Handbook of Fuzzy Sets Series*, vol. 4, *Fuzzy Sets in Decision Analysis, Operations Research and Statistics*. Kluwer Academic Press. pp. 31–68.
- ISO/IEC 9126 (1991). *Information Technology – Software Product Evaluation – Quality Characteristics and Guidelines for their Use*. First edition, 1991-12-15, reference number ISO/IEC 9126: 1991(E).
- Klement, E.P., R. Mesiar and E. Pap (2002). Triangular norms. Position paper I: Basic analytical and algebraic properties. *Technical Report FLLL-TR-0208*, Johannes Kepler Universität Linz, Fuzzy Logic Laboratorium Linz-Hagenberg.
http://www.flll.jku.at/research/technical_reports/flll-tr-0208.pdf
- Kolmogorov, A. (1930). Sur la notion de moyenne. *Atti delle Reale Accademia Nazionale dei Lincei Mem. Cl. Sci. Mat. Mnatur. Sez. 12*, pp. 323–343 (in French).
- Mayor, G., and E. Trillas (1986). On the representation of some Aggregation Functions. In *Proceedings of the XVth IEEE-International Symposium on Multiple-Valued Logic*. pp. 110–114.
- Mesiar, R., and M. Komorníková. (1997). Aggregation operators. In *Proceeding of the XI Conference on Applied Mathematics (PRIM'96)*. Novi Sad. pp. 193–211.
- Moraga, C. (2001). Neuro-fuzzy modeling between minimum and maximum. In *Proceeding of the Workshop on Computational Intelligence, Theory and Applications*. Press Technical University of Niš, Yugoslavia.
- Ovchinnikov, S. (1998). On robust aggregation procedures. In B. Bouchon-Meunier (Ed.), *Aggregation and Fusion of Imperfect Information*. Physic Verlag. pp. 3–10.
- Pfeiffer, P.E. (1979). *Concepts of Probability Theory*. Dover Publications, 2nd edition.
- Schweizer, B., and A. Sklar (1960). Statistical metric spaces. *Pacific J. Math.*, **10**, 313–334.
- Schweizer, B., and A. Sklar (1983). *Probabilistic Metric Spaces*. North Holland.
- Silvert, W. (1979). Symmetric summation: a class of operations on fuzzy sets. *IEE Transactions on Systems, Man and Cybernetics*, **9**, 659–667.
- Sugeno, M. (1974). *Theory of Fuzzy Integrals and its Application*. Phd. thesis. Tokyo Institute of Technology.
- Yager, R.R. (1988). On ordered weighted averaging aggregation operators in multi-criteria decision-making. *IEE Transactions on Systems, Man and Cybernetics*, **18**, 183–190.
- Zimmermann, H.J., and P. Zysno (1980). Latent connectives in human decision making. *Fuzzy Sets and Systems*, **4**, 37–51.

A. Čaplinskas is a principal researcher at the Institute of Mathematics and Informatics. The area of his scientific interest includes software engineering, information system engineering, legislative engineering, and knowledge-based systems.

J. Gasperovič is a doctoral student at the Institute of Mathematics and Informatics. Her research area encompasses formal, semi-formal and lightweight formal specification languages and their application in software engineering and information systems engineering.

Informacinių sistemų specifikavimo kalbų vidinės kokybės charakteristikų agregavimas

Albertas ČAPLINSKAS, Jelena GASPEROVIČ

Straipsnyje nagrinėjamas informacinių sistemų specifikavimo kalbų vidinės kokybės charakteristikų agregavimo uždavinys. Trumpai apžvelgiama agregavimo teorija. Parodyta, kad tokios charakteristikos gali būti siejamos keturių tipų sąryšiais ir pasiūlyta, kaip konstruoti atitinkamas t -normas ir t -konormas kokybės charakteristikoms agreguoti kiekvienu iš šių keturių atvejų. Pasiūlyta euristika, taikytina agreguojant charakteristikas, kurių reikšmės gautos panaudojus nevissai adekvačius matavimo būdus arba šiek tiek netikslias matavimo procedūras.