Interval-Based Parameter Recognition with the Trends in Multiple Estimations

V. Terziyan, S. Puuronen, H. Kaykova

University of Jyvaskyla, P.O.Box 35, FIN-40351 Jyvaskyla, Finland
phone: +358 14 603 260, fax: +358 14 603 011,
e-mail: {vagan, sepi, helen}@jytko.jyu.fi

Abstract

This paper considers the context sensitive approach to handle interval knowledge acquired from multiple knowledge sources. Each source gives its estimation of the value of some parameter *x*. The goal is to process all the intervals in a context of trends caused by some noise and derive resulting estimation that is more precise than the original ones and also takes into account the context noise. The main assumption used is that if a knowledge source guarantees smaller measurement error (estimated interval is shorter) then this source in the same time is more resistant against the effect of noise. This assumption allows us to derive and process trends among intervals and end up to shorter resulting estimated interval than any of the original ones.

1 Introduction

It is generally accepted that knowledge has a contextual component. Acquisition, representation, and exploitation of knowledge in context would have a major contribution in knowledge representation, knowledge acquisition, and explanation [4].

It is noticed in [5] that knowledge-based systems do not use correctly their knowledge. Knowledge being acquired from human experts does not usually include its context.

Contextual component of knowledge is closely connected with eliciting expertise from one or more experts in order to construct a single knowledge base (or, for example as in [3], for co-operative building of explanations). If more than one expert available, one must either select the opinion of the best expert or pool the experts' judgements [28]. It is assumed here that when experts' judgements are pooled, collectively they offer sufficient cues leading to smaller uncertainty.

In recognition of some pattern it is also possible to handle context of recognition using one of two decontextualization techniques: a) using the same "recognizer" in different contexts and then combining recognition results (as it is shown in Fig. 1); b) using different "recognizers" in the same context.

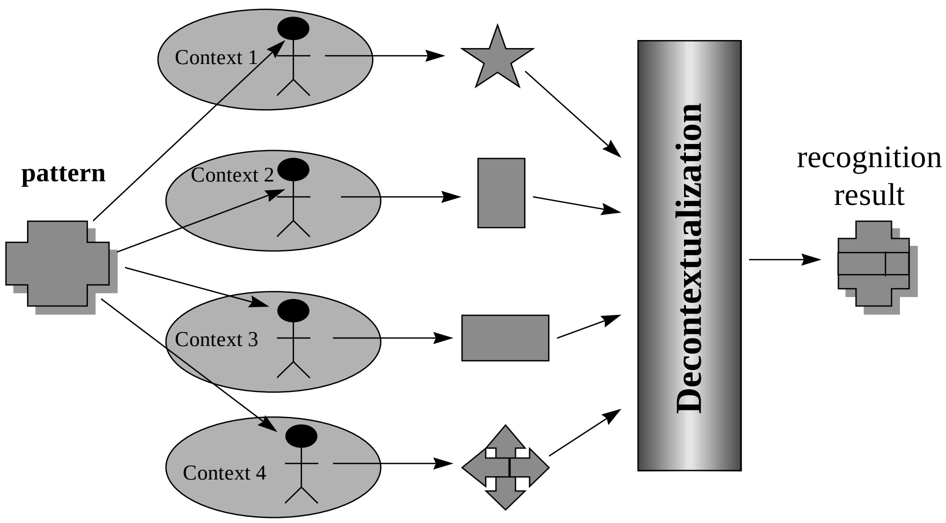

Fig. 1: Context in pattern recognition

All information about the real word comes from two sources: from measurements, and from experts [16]. Measurements are not absolutely accurate. Every measurement instrument usually has the guaranteed upper bound of the measurement error. The measurement result is expected to lie in the interval around the actual value. This inaccuracy leads to the need to estimate the resulting inaccuracy of data processing. When experts are used to estimate the value of some parameter, intervals are commonly used to describe degrees of belief [23]. Experts are often uncertain about their degrees of belief making far larger estimation errors than the boundaries accepted by them as feasible [12]. In both cases we deal with interval uncertainty, i.e. we do not know exact values of parameters, only intervals where the values of these parameters belong to. A number of methods to define operations on intervals that produce guaranteed precision have been developed in [21], [22], [18], and [1] among others.

In many real life cases there is also some noise which does not allow direct measurement of parameters. To get rid of this noise it is necessary to subtract its value from the result of measurement. The noise can be considered as an undesirable effect to the evaluation of a parameter in the context. The subtraction of the noise in this sense has certain analogy with the decontextualization [20], [13], [7]. When effect of noise is not known the help of decontextualization using several coexisting experts (recognizers, knowledge sources) might estimate it (Fig. 2).

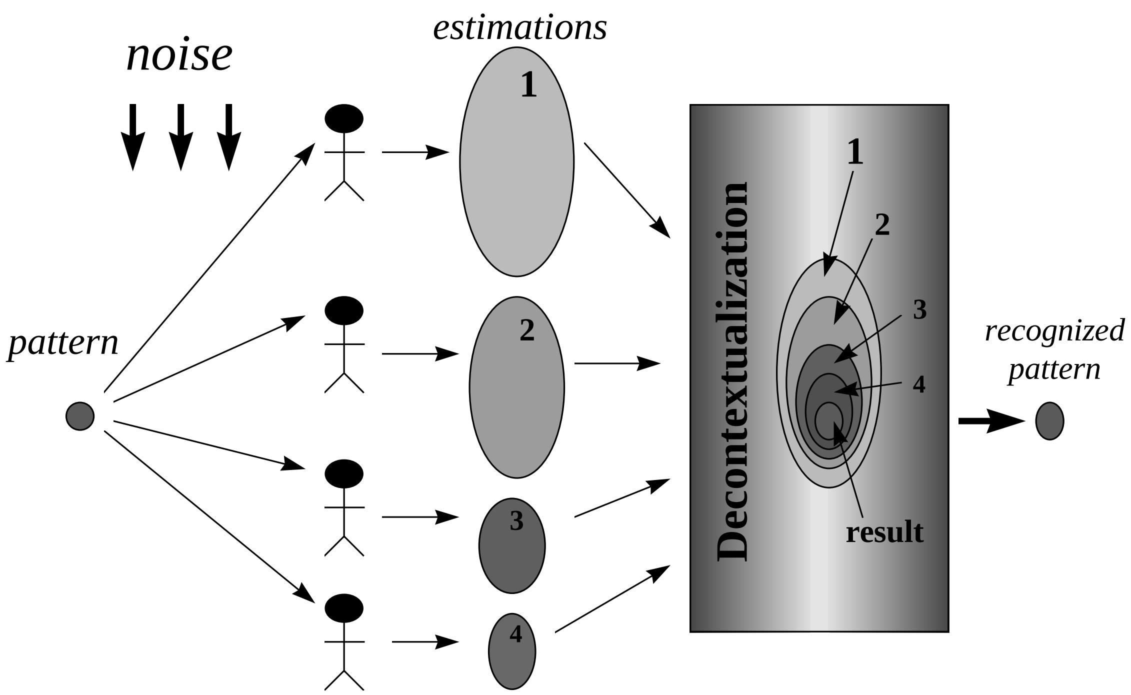

Fig. 2: Decontextualization of noise in pattern recognition

Some geometrical heuristics were used in [6] to solve this problem without enough mathematical justification. It is natural to assume that different measurement instruments as well as different experts possess different resistance against the influence of noise. Using measurements from several different instruments as well as estimations from multiple experts we try to discover the effect caused by noise and thus be able to derive the decontextualized measurement result like it is shown in the example in Fig.2.

This paper considers a context sensitive approach to handle interval knowledge acquired from multiple knowledge sources. Each source is assumed to give its evaluation, i.e. an estimated interval to which the value of a parameter *x* belongs. The goal is to process all the given intervals in the contexts of trends and derive more precise estimation of the value of parameter from them. The quality of each source is considered from two points of view: first, the value of guaranteed upper bound of measurement error, and second, the value of a resistance against a noise. These are assumed to occur together.

The main assumption in this paper is as follows. The estimation of some parameter *x* given by more accurate knowledge source (*i.e. source guarantees smaller upper bound of measurement error*) is supposed to be closer to the actual value of parameter *x* (*i.e. source is more resistant against a noise of estimation*). The assumption allows us to derive different trends in cases when there are multiple estimation that result to shorter estimation intervals.

In chapter 2 we present our decontextualization process and some of it main characteristics in the case of one trend. Next chapter discusses about one way to formulate groups of trends and its relation to decontextualization process. Chapter 4 discusses combining results of several trends into one resulting interval. The last chapter includes very short conclusion.

2 Decontextualization

In this chapter we consider a decontextualization process that is used to improve interval estimation by processing recursively more bounded intervals against less bounded ones.

Let there be *n* knowledge sources (human beings or measurement instruments) which are asked to make estimations of the value of a parameter *x*. Each knowledge source *i*, $i=1,\ldots,n$ gives his estimation $L_{[a_i,b_i]}$, $a_i < b_i$ as a closed interval into which he is sure that the value of the parameter belongs to. *L* is the estimation predicate as follows:

Definition 2.1:

The *range* of a parameter *x* is the length $b_0 - a_0$ of the interval from the estimation $L_{[a_0,b_0]}$, which includes all possible interval estimations $L_{[a_i,b_i]}$, $i=1,\ldots,n$ of this parameter.

Let us assume that all the knowledge sources are effected by the same misleading noise in the context of estimation. Different knowledge sources are effected by such a noise in a different way. The main assumption used in this paper is that: if a knowledge source guarantees smaller measurement error (interval estimation is more narrow), then this source is also more resistant against the effect of noise. This assumption also means that the estimated value given by more precise knowledge source is supposed to be closer to the actual value of the parameter *x*. This assumption is used when we derive trends of intervals towards the actual value of the parameter *x*.

The process advances decontextualization of an interval from the context of another interval (the step of decontextualization process) by pairs beginning from the shortest and second shortest intervals. The next step of decontextualization process is made decontextualization of the resulting interval of the first step from the context of the third shortest original interval. This is continued until all original intervals have been participated the process. The result of the last step is the result of the whole decontextualization process.

Definition 2.2:

The *uncertainty* u_i of interval estimation $L_{[a_i,b_i]}$ is equal to the length of the interval:

$$u_i = b_i - a_i,\ i=1,\ldots,n.$$

To be precise it is necessary to mention that in a general case the value of uncertainty should be standardized with the range of a parameter estimated, like the following:

$$L_{[a_i,b_i]} = \begin{cases} 1, & \text{if the parameter is estimated} \\ & \text{to belong to the interval } [a,b]; \\ 0, & \text{otherwise.} \end{cases}$$

$$u_i^{st} = \frac{b_i - a_i}{b_0 - a_0}, \; i=1,\dots,n.$$

In this paper, however, we use and compare different estimations of the same parameter within the same range. That is why it is not essential to standardize a value of uncertainty and we can use the Definition 2.2 working with uncertainty.

Definition 2.3:

The *quality* q_i of interval estimation $L_{[a_i, b_i]}$ is the reverse of its uncertainty, i.e.:

$$q_i = \frac{1}{u_i}, \; i=1,\dots,n.$$

2.1 Operating with two intervals

Definition 2.4:

The result of one step of the decontextualization process with two interval estimations $L_{[a_i, b_i]}$ and $L_{[a_j, b_j]}$, $u_i \neq u_j$, $i=1,\dots,n$ is:

$$L_{[a_i,b_i]}^{L_{[a_j,b_j]}} = L_{\left[a_i + \frac{u_i^2 \cdot (a_i - a_j)}{u_j^2 - u_i^2}, \; b_i + \frac{u_i^2 \cdot (b_i - b_j)}{u_j^2 - u_i^2}\right]}.$$

The above formulas for calculating the resulting interval was selected because they satisfy three main requirements:

- the resulting interval should be shorter than the original ones,
- the longer the original intervals are the longer should the resulting interval be, and
- shorter of the two intervals should locate closer to the resulting interval than the longer one.

In the following we will prove that the selected formulas fulfill these three main requirements.

The following theorem defines the relationships between the uncertainties of the original and the resulting interval estimations.

Theorem 2.1:

Let it be that:

$$L_{[a_i,b_i]}^{L_{[a_j,b_j]}} = L_{[a_{res}, \; b_{res}]},$$

where a_{res} and b_{res} are as in the right hand part of the Definition 2.4.

Then:

a) $u_{res} = \frac{u_i \cdot u_j}{u_i + u_j}$, b) $u_{res} < u_i$,

c) $u_{res} < u_j$, d) $q_{res} = q_i + q_j$.

Proof:

a) According to the Definition 2.4:

$$a_{res} = a_i + \frac{u_i^2 \cdot (a_i - a_j)}{u_j^2 - u_i^2} \quad \text{and}$$

$$b_{res} = b_i + \frac{u_i^2 \cdot (b_i - b_j)}{u_j^2 - u_i^2}.$$

Definition 2.2 gives us that:

$$u_{res} = b_{res} - a_{res} = (b_i - a_i) + \frac{u_i^2}{u_j^2 - u_i^2} \cdot$$

$$\cdot ((b_i - a_i) - (b_j - a_j)) =$$

$$= u_i + \frac{u_i^2}{u_j^2 - u_i^2} \cdot (u_i - u_j) =$$

$$= u_i - \frac{u_i^2}{u_i + u_j} = \frac{u_i \cdot u_j}{u_i + u_j};$$

Thus: $$u_{res} = \frac{u_i \cdot u_j}{u_i + u_j};$$

b) Let us suppose that: $u_{res} \geq u_i$,

then according to (a) we receive:

$$\frac{u_i \cdot u_j}{u_i + u_j} \geq u_i \Rightarrow \frac{u_i + u_j}{u_j} \leq 1 \Rightarrow u_i \leq 0,$$

which contradicts the Definition 2.2. Thus: $u_{res} < u_i$;

c) Prove is similar as for (b).

d) From the Definition 2.3 it results that:

$$u_i = \frac{1}{q_i}, \; i=1,\dots,n.$$

Applying (a) we receive that:

$$\frac{1}{q_{res}} = \frac{\frac{1}{q_i}\cdot\frac{1}{q_j}}{\frac{1}{q_i}+\frac{1}{q_j}} = \frac{1}{q_i+q_j};$$

Thus: $$q_{res} = q_i + q_j.$$

Theorem 2.2:

Let it be that: $L_{[a_i,b_i]}^{L_{[a_j,b_j]}} = L_{[a_{res_1},\ b_{res_1}]}$, and $L_{[a_i,b_i]}^{L_{[a_k,b_k]}} = L_{[a_{res_2},\ b_{res_2}]}$, where a_{res_1}, a_{res_2}, b_{res_1}, and b_{res_2} are as in the right hand part of Definition 2.4. Let it be that: $u_j < u_k$. Then: $u_{res_1} < u_{res_2}$.

Proof:

$$u_j < u_k \Rightarrow u_i u_j < u_i u_k \Rightarrow$$

$$\Rightarrow u_i u_j + u_j u_k < u_i u_k + u_j u_k \Rightarrow$$

$$\Rightarrow (u_i + u_k) u_j < (u_i + u_j) u_k \Rightarrow \frac{u_j}{u_i + u_j} < \frac{u_k}{u_i + u_k} \Rightarrow$$

$$\Rightarrow \frac{u_i \cdot u_j}{u_i + u_j} < \frac{u_i \cdot u_k}{u_i + u_k} \Rightarrow u_{res_1} < u_{res_2}.$$

Theorem 2.3:

Let it be that: $L_{[a_j,b_j]}^{L_{[a_i,b_i]}} = L_{[a_{res_1},b_{res_1}]}$, and $L_{[a_k,b_k]}^{L_{[a_i,b_i]}} = L_{[a_{res_2},\ b_{res_2}]}$, where a_{res_1}, a_{res_2}, b_{res_1}, and b_{res_2} are as in the right hand part of the Definition 2.4. Let it be that $u_j < u_k$. Then: $u_{res_1} < u_{res_2}$.

Proof: Similarly as Theorem 2.2.

It is easy to see that the formula of the resulting uncertainty calculation has very simple physical interpretation. It is shown in Fig. 3.

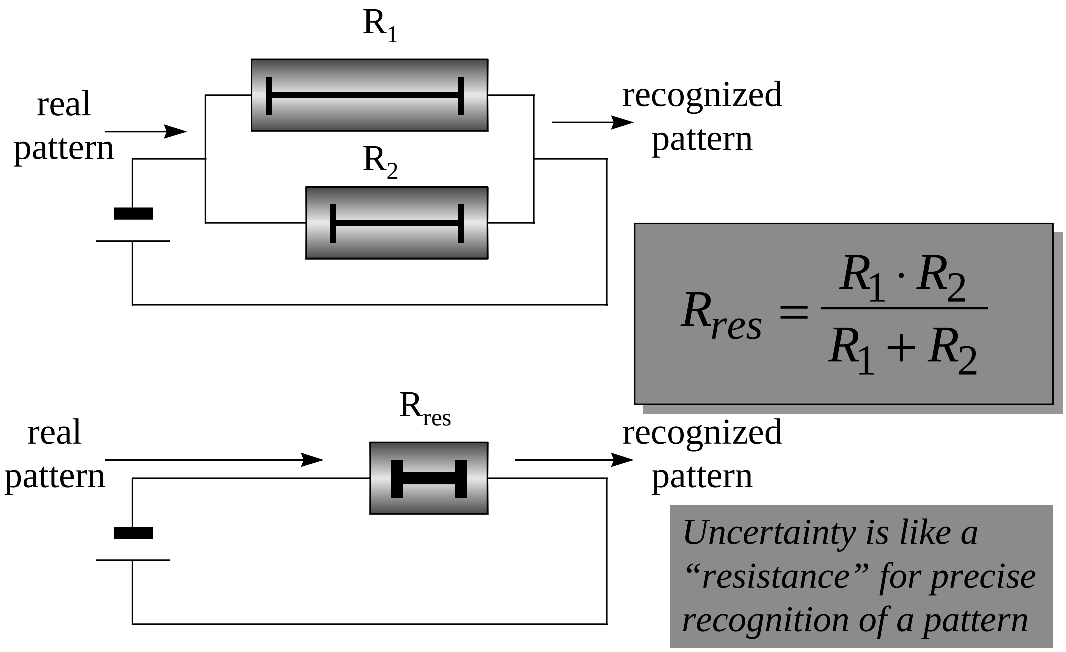

Fig. 3: Physical interpretation of decontextualization

Every interval uncertainty can be considered as certain "resistance" to exact recognition of a pattern. However if we use several estimations and connect the appropriate resistance in a parallel scheme, then the resulting resistance (according well known from physics formula, Fig. 3) will be smaller then every separate one.

Another possible interpretation of the step of decontextualization formula is based on an extrapolation of the decontextualized value using interval functions. Extrapolation is based on assumption of linearity of these functions within one step of decontextualization.

The two linear functions are considered: $a = f(u)$, that connects points $(u_i, a_i), (u_j, a_j)$, and $b = \varphi(u)$ that connects points $(u_i, b_i), (u_j, b_j)$ as shown in Fig. 4.

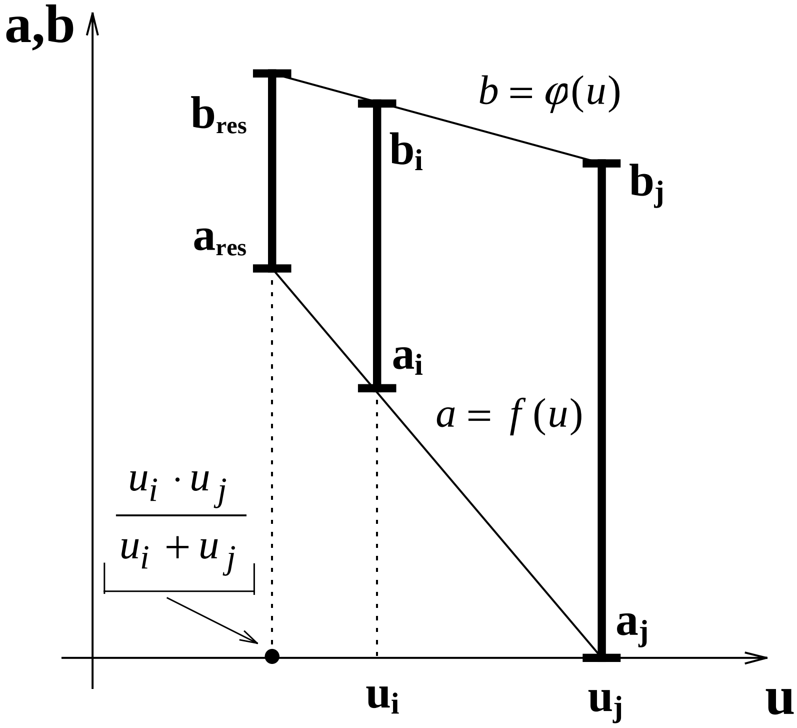

Fig. 4: Deriving decontextualized interval by linear extrapolation

To obtain $f\left(\frac{u_i \cdot u_j}{u_i + u_j}\right)$ and $\varphi\left(\frac{u_i \cdot u_j}{u_i + u_j}\right)$ values we should solve the equations:

$$\frac{u_i - \frac{u_i \cdot u_j}{u_i + u_j}}{u_j - u_i} = \frac{a_i - a_{res}}{a_j - a_i};$$

$$\frac{u_i - \frac{u_i \cdot u_j}{u_i + u_j}}{u_j - u_i} = \frac{b_i - b_{res}}{b_j - b_i}.$$

The solution of these equations gives us the following values:

$$a_{res} = a_i + \frac{u_i^2 \cdot (a_i - a_j)}{u_j^2 - u_i^2}; \quad b_{res} = b_i + \frac{u_i^2 \cdot (b_i - b_j)}{u_j^2 - u_i^2},$$

which is exactly the same as if we make one step of decontextualization process accordingly to the Definition 2.4. Also this fact motivates the selection of the formula in the Definition 2.4 from the point of view of the possibility to obtain result of decontextualization using linear extrapolation.

Theorem 2.4:

$$L_{[a_i, b_i]}^{L_{[a_j, b_j]}} = L_{[a_j, b_j]}^{L_{[a_i, b_i]}}.$$

Proof:

$$L_{[a_i, b_i]}^{L_{[a_j, b_j]}} = L_{\left[a_i + \frac{u_i^2 \cdot (a_i - a_j)}{u_j^2 - u_i^2},\ b_i + \frac{u_i^2 \cdot (b_i - b_j)}{u_j^2 - u_i^2}\right]} =$$

$$L\left[\frac{u_j^2\cdot a_i-u_i^2\cdot a_j}{u_j^2-u_i^2},\frac{u_j^2\cdot b_i-u_i^2\cdot b_j}{u_j^2-u_i^2}\right]=$$

$$L\left[\frac{u_i^2\cdot a_j-u_j^2\cdot a_i}{u_i^2-u_j^2},\frac{u_i^2\cdot b_j-u_j^2\cdot b_i}{u_i^2-u_j^2}\right]=$$

$$L\left[a_j+\frac{u_j^2\cdot(a_j-a_i)}{u_i^2-u_j^2},\ b_j+\frac{u_j^2\cdot(b_j-b_i)}{u_i^2-u_j^2}\right]=L^{L[a_i,b_i]}_{[a_j,b_j]}.$$

2.2 Measuring distance between intervals

To prove that the resulting interval is located closer to smallest of the original two intervals, we need to define distance function between intervals.

There are many approaches to define distance between any two entities (attributes, terms) based on their numerical or semantic closeness. For example the semantic closeness between terms is a measure of how closely terms are related in the classification schema [27].

Distance metric used by Rada et al. [25] represents the conceptual distance between concepts. Rada et al. uses only the path length to determine this conceptual distance, with no consideration of node or link characteristics. Distance is measured as the length of the path representing the traversal from the first classification term to the second. The closeness of terms ranges from 1 (identical terms) to 0 (which represents that terms are not semantically close, although it does not mean that they are disjoint in the classification schema).

Rocha [26] has suggested a method to “fuzzify” conversation theory, by calculating continuously varying conceptual distances between nodes in an entailment mesh, on the basis of the number of linked nodes they share.

In psychology, rudimentary associative networks have been created through experiments in which subjects were given a word (say cat), and were asked which other word first came to mind (e.g. dog, mouse, or milk). The more often a certain word b is given in response to the cue word a, the stronger the association from a to b. Since this approach usually only finds a small number of associations for any given word, association strengths for links between other words are calculated by taking into account indirect associations (e.g. knowing the strengths of dog -> cat and cat -> mouse would allow one to calculate the strength of dog -> mouse). Note that such associations are in general asymmetric. For example, when cued with penguin the probability that you would say bird is not so small, whereas the probability to respond with penguin, when cued with bird is virtually zero. This methodology, however, requires a lot of work from designers and users, and is only useful for simple, well-known items like common words [14].

Heylighen prefers the metaphor of bootstrapping [14]. The problem with correspondence epistemologies is that they lack grounding: everything is built on top of the symbols, which constitute the atoms of meaning; yet, the symbols themselves are not supported. The advantage of a coherence epistemology is that there is no need for a fixed ground or foundation on which to build models: coherence is a two-way relation. In other words, coherent concepts support each other. The dynamic equivalent of this mutual support relation may be called "bootstrapping": Model A can be used to help construct model B, while B is used to help construct A. It is as if I am pulling myself up by my own bootstraps: while my arms (A) pull up my boots (B), my boots at the same time - through my legs, back and shoulders - push up my arms. The net effect is that more (complexity, meaning, quality, ...) is produced out of less. This is the hallmark of self-organization: the creation of structure without need for external intervention.

Associative networks are in principle more general and more flexible, allowing the expression of different "fuzzy", "intuitive" or even "subconscious" relations between concepts. Such networks have been regularly suggested as models of how the brain works. They are similar to the presently popular "neural" networks, except that the latter are typically used as directed, information processing systems, which are given a certain pattern of stimuli as input and are supposed to produce the correct response to that pattern as output. In the present "bootstrapping" perspective, there is no overall direction or sequence leading from inputs to outputs; there are only nodes linked to each other by associations, in such a way that they are coherent with each other and with the user's understanding of the knowledge domain. Associative networks could be created by the same type of knowledge elicitation techniques, where a user enters a number of concepts and links and is prompted by the system to add further links and concepts under the main constraint of avoiding ambiguity. These links must then be attributed some variable degree of strength. However, the very weak requirement of "associativity" allows virtually any pair of concepts to be linked, if only with a very small link strength [14].

In order to measure the distance between two concepts in a mind, Jorgensen measures a distance between two concepts, which he calls *psy* [15]. It has been suggested to assign an arbitrary distance of *n* units to the separation between two concepts such as "Concept *A*" and "Concept *B*" and then ask a subject to tell us how far other concepts (*C* and *D*) are from each other in these units. One assumption used is that our given distance, *n* units, is a close enough to the distance one wishes to measure so as to avoid the errors which are typical of people trying to deal with very small or very large numbers. The basic premise that the relationship of concepts in a mind can and should be measured as a ratio of a relationship between two concepts in a mind is as valid as measuring the distance between two fixed points on a rigid body as a ratio of the distance between two given points on that body. Jorgensen presents one interactive measuring instrument that uses physical distances on a computer screen (or more accurately, the visual angles that is subtended by pairs of points on a screen) rather than numbers. He introduces an appropriate Java program as an interactive visual instrument for measuring the distance between concepts. Two benefits have been distinguished: reducing or eliminating the subliminal baggage that numbers carry along with them, and eliminating some of the artifacts that numerical scales can introduce due to the different ways different people think about numbers.

Lynch and Chen represent the information retrieval problem as a search task, where the goal was to identify the most relevant descriptors associated with the searcher's search terms [19]. The costs to each activated path on the semantic network have been assigned that were based on the nodes visited, the types of links traversed, and the number of links in the path. Cost was a metric that is used to indicate the "semantic distance" (relevance) of terms. A branch-and-bound algorithm guides the search. This algorithm computed and ordered costs for each partial path and expanded the least-cost path. It terminated when all relevant paths were explored or when the algorithm ran out of activation levels (2-links activation for each source term). This algorithm allows the system to identify most relevant concepts in a large network of knowledge. The algorithm acts as the system control module: selects appropriate knowledge sources, activates nodes and links, calculates scores of relevance, and suggests to users the most relevant topics, concepts, or descriptors.

Access to and navigation through multimedia material is usually a system that needs multiple dimensions of classification. These classification schemas are often specialized, and therefore a more general user may not be familiar with the organization of the classification schema or the terms that are employed. Supporting all access to the material through classification schema raises a number of issues. Tailor and Tudhope [27] has presented a hypermedia architecture that is supported by classification schema. Semantic closeness measures have been developed to measure the closeness of terms in the schema which provides a platform for high-level navigation tools, which can provide flexible access tools to a collection of material. Two higher level navigation tools, navigation via media similarity and best-fit generalization, have been developed. The similarity coefficients are extended in that similarity is judged on the "semantic closeness" of the sets of classification terms that are attached to the media nodes. The similarity coefficient therefore needs to be able to handle sets of classification terms with varying lengths, with non-exact matches of terms, where the pairing of terms between media nodes may not be immediately obvious.

Brooks reports two experiments that investigated the semantic distance model (SDM) of relevance assessment [6]. In the first experiment graduate students of mathematics and economics assessed the relevance relationships between bibliographic records and hierarchies of terms composed of classification headings or help-menu terms. The relevance assessments of the classification headings, but not the help-menu terms, exhibited both a semantic distance effect and a semantic direction effect as predicted by the SDM. Topical subject expertise enhanced both these effects. The second experiment investigated whether the poor performance of the help-menu terms was an experimental design artifact reflecting the comparison of terse help terms with verbose classification headings. In the second experiment the help-menu terms were compared to a hierarchy of single-word terms where they exhibited both a semantic distance and semantic direction effect.

Foo at al [10] propose and define a modification of Sowa's metric on conceptual graphs. The metric is computed by locating the least subtype which subsumes the two given types, and adding the distance from each given type to the subsuming type.

Cugini at al [9] uses the nearest neighbor paradigm as a heuristic to get semantically similar documents to cluster in the same spatial region. Each document has a "position" in semantic space, represented as a vector of keyword strengths. Since there are at most three spatial dimensions for visualization but potentially many more keywords, they cannot simply map documents directly from "keyword space" to geometric space. Instead, they try to find a linear order that keeps semantically close documents if not adjacent, then at least nearby. They used a simple nearest neighbor algorithm to order the documents: given some initial choice, each document in the sequence is the nearest (of those not already in the sequence) neighbor to its predecessor. The semantic distance between any two documents is based on each document's keyword strength vector and can be computed in two ways: simple Euclidean distance, or as the angle between the vectors.

Statistical methods can be used to analyze database contents in an attempt to group items according to some measure of their semantic closeness. For example the contents of a document store could be grouped corresponding to matching keywords. Analysis performed on these information stores typically result in a number of 'scores' for documents, which can then be used to create a suitable mapping into Benediktine space. The closer objects are semantically then the closer they will be within the data environment. Systems, which adopt this approach, include VIBE [24] and BEAD [8], though the original idea of VIBE has been developed further and extended into three dimensions to produce VR-VIBE [2].

Vineta is a prototype information visualization system developed by Uwe Krohn [17]. Vineta allows the visualization, browsing and querying of large bibliographic data without resorting to typing and revising keyword based queries. Similar to VR-VIBE and BEAD visualizations, Vineta presents documents and terms as graphical objects within a three dimensional space, the navigation space. The positioning of these objects within that space encodes the semantic relevance between documents, terms and the user's interests. Vineta is built upon the premise that navigation through an information space can be an effective means of retrieving information of interest. Krohn states that informational navigation is strongly connected with the human intuitive comprehension of abstract facts by means of analogies with familiar concepts such as location or motion. Vineta uses spatial proximity to represent semantic similarity between objects (i.e. documents).

Instance-based learning techniques typically handle continuous and linear input values well, but often do not handle nominal input attributes appropriately. The Value Difference Metric (VDM) was designed by Wilson and Martinez [29] to find reasonable distance values between nominal attribute values, but it largely ignores continuous attributes, requiring discretization to map continuous values into nominal values. Wilson and Martinez propose new heterogeneous distance functions, called the Heterogeneous Value Difference Metric (HVDM), the Interpolated Value Difference Metric (IVDM), and the Windowed Value Difference Metric (WVDM). These new distance functions are designed to handle applications with nominal attributes, continuous attributes, or both. In experiments on 48 applications the new distance metrics achieve higher classification accuracy on average than three previous distance functions on those datasets that have both nominal and continuous attributes.

As it was mentioned in the Wilson and Martinez review [29] there are many learning systems that depend upon a good distance function to be successful. A variety of distance functions are available for such uses, including the Minkowsky, Mahalanobis, Camberra, Chebychev, Quadratic, Correlation, and Chi-square distance metrics; the Context-Similarity measure; the Contrast Model; hyperrectangle distance functions and others. Several of these functions are defined in Fig. 5.

Minkowsky:

$$D(x,y)=\left(\sum_{i=1}^{m}|x_i-y_i|^r\right)^{1/r}$$

Euclidian:

$$D(x,y)=\sqrt{\sum_{i=1}^{m}(x_i-y_i)^2}$$

Camberra:

$$D(x,y)=\sum_{i=1}^{m}\frac{|x_i-y_i|}{|x_i+y_i|}$$

Manhattan/city block:

$$D(x,y)=\sum_{i=1}^{m}|x_i-y_i|$$

Quadratic:

$$D(x,y)=(x-y)^T\cdot Q(x-y)=$$

$$=\sum_{j=1}^{m}\left(\sum_{i=1}^{m}(x_i-y_i)\cdot q_{ji}\right)\cdot(x_j-y_j)$$

Q is a problem-specific positive definite mxm weight matrix

Chebychev:

$$D(x,y)=\max_{i=1}^{m}|x_i-y_i|$$

Mahalanobis:

$$D(x,y)=[\det V]^{1/m}(x-y)^T\cdot V^{-1}(x-y)$$

V is the covariance matrix of $A_1\cdot A_m$, and A_j is the vector of values for attribute j occurring in the training set instances 1…n.

Correlation:

$$D(x,y)=\frac{\sum_{i=1}^{m}(x_i-\bar{x}_i)\cdot(y_i-\bar{y}_i)}{\sqrt{\sum_{i=1}^{m}(x_i-\bar{x}_i)^2\cdot\sum_{i=1}^{m}(y_i-\bar{y}_i)^2}}$$

$\bar{x}_i=\bar{y}_i$ and is the average value for attribute i occurring in the training set.

Chi-square:

$$D(x,y)=\sum_{i=1}^{m}\frac{1}{sum_i}\cdot\left(\frac{x_i}{size_x}-\frac{y_i}{size_y}\right)^2$$

sum_i is the sum of all values for attribute i occurring in the training set, and $size_x$ is the sum of all values in the vector x.

Kendall's Rank Correlation:

$$D(x,y)=1-\frac{2}{n\cdot(n-1)}\cdot$$

$$\cdot\sum_{i=1}^{m}\sum_{j=1}^{i-1}sign(x_i-x_j)\cdot sign(y_i-y_j),$$

$$sign(x)=\begin{cases}-1, & x<0;\\ 0, & x=0;\\ 1, & x>0.\end{cases}$$

Fig. 5: Equations of selected distance functions according to Wilson and Martinez review [29]; x and y are vectors of m attribute values.

An interval can be considered as a vector with two attributes, which are the values of its endpoints. For measuring distance between two intervals we select a variation of Chebyshev distance function (Fig. 5).

Definition 2.5:

Let us have two interval estimations $L_{[a_i,b_i]}$ and $L_{[a_j,b_j]}$, $i,j=1,\dots,n$.

The *distance* between these opinions is as follows:

$$D(L_{[a_i,b_i]},L_{[a_j,b_j]})=\max(abs(a_j-a_i),abs(b_j-b_i)).$$

Theorem 2.5:

If it holds that: $L_{[a_i,b_i]}^{L_{[a_j,b_j]}}=L_{[a_{res},\ b_{res}]}$,

and $u_i<u_j$, then:

$$D(L_{[a_{res},b_{res}]},L_{[a_i,b_i]})<D(L_{[a_{res},b_{res}]},L_{[a_j,b_j]}).$$

Proof:

Using Definitions 2.4 and 2.5 we receive:

$$(D(L_{[a_{res},b_{res}]},L_{[a_i,b_i]})=$$
$$=\max(abs(a_i-a_{res}),abs(b_i-b_{res}))=$$
$$=\max(abs(\frac{u_i^2\cdot(a_j-a_i)}{u_j^2-u_i^2}),abs(\frac{u_i^2\cdot(b_j-b_i)}{u_j^2-u_i^2})))<$$
$$<(\max(abs(\frac{u_j^2\cdot(a_j-a_i)}{u_j^2-u_i^2}),abs(\frac{u_j^2\cdot(b_j-b_i)}{u_j^2-u_i^2}))=$$
$$=\max(abs((a_j-a_i)-\frac{u_i^2\cdot(a_i-a_j)}{u_j^2-u_i^2}),abs((b_j-$$
$$-b_i)-\frac{u_j^2\cdot(b_i-b_j)}{u_j^2-u_i^2}))=\max(abs(a_j-$$
$$-a_{res}),abs(b_j-b_{res}))=D(L_{[a_{res},b_{res}]},L_{[a_j,b_j]})).$$

2.3 Operating with several intervals

The process of decontextualization with several intervals was described in the beginning of this chapter. We describe now this step-by-step process formally.

Let there be n interval estimations $L_{[a_i,b_i]}$, $i=1,\dots,n$, $u_i<u_{i+1}$, $i=1,\dots,n-1$, $n\geq 2$. The resulting interval estimation:

$$L_{[a_{res},b_{res}]}=L_{[a_1,b_1]}^{L_{[a_2,b_2]}^{L_{\dots}^{L_{[a_n,b_n]}}}}$$

can be calculated recursively as follows:

$$L_{[a_{res_1},b_{res_1}]}=L_{[a_1,b_1]};$$

$$L_{[a_{res_i},b_{res_i}]}=L_{[a_{res_{i-1}},b_{res_{i-1}}]}^{L_{[a_i,b_i]}},i=2,\dots,n;$$

$$L_{[a_{res},b_{res}]}=L_{[a_{res_n},b_{res_n}]}.$$

Theorem 2.6:

If it holds that:

$(L_{[a_i,b_i]}^{L_{[a_j,b_j]}})^{L_{[a_k,b_k]}}=L_{[a_{res'},\ b_{res'}]}$ and $L_{[a_i,b_i]}^{(L_{[a_j,b_j]}^{L_{[a_k,b_k]}})}=L_{[a_{res''},\ b_{res''}]}$, then: $u_{res'}=u_{res''}$.

Proof:

$$u_{res'}=\frac{\left(\frac{u_i\cdot u_j}{u_i+u_j}\right)\cdot u_k}{\left(\frac{u_i\cdot u_j}{u_i+u_j}\right)+u_k}=\frac{\frac{u_i\cdot u_j\cdot u_k}{u_i+u_j}}{\frac{u_i\cdot u_j+u_i\cdot u_k+u_j\cdot u_k}{u_i+u_j}}=$$

$$=\frac{u_i\cdot u_j\cdot u_k}{u_i\cdot(u_j+u_k)+u_j\cdot u_k}=$$

$$=\frac{\frac{u_i\cdot u_j\cdot u_k}{u_j+u_k}}{\frac{u_i\cdot(u_j+u_k)+u_j\cdot u_k}{u_j+u_k}}=$$

$$\frac{u_i\cdot\left(\frac{u_j\cdot u_k}{u_j+u_k}\right)}{u_i+\left(\frac{u_j\cdot u_k}{u_j+u_k}\right)}=u_{res''}.$$

An example: Let us suppose that three knowledge sources 1, 2, and 3 evaluate the value of the attribute x to be in the following intervals:

$$L_{[a_1,b_1]} = L_{[9,12]},\ L_{[a_2,b_2]} = L_{[6,11]},$$

$$L_{[a_3,b_3]} = L_{[0,10]}.$$

The above intervals are already in ascending order according to their uncertainties. The resulting interval is derived by the recursive procedure above:

$$L_{[a_{res_1},b_{res_1}]} = L_{[a_1,b_1]} = L_{[9,12]};$$

$$L_{[a_{res_2},b_{res_2}]} = L_{[a_{res_1},b_{res_1}]}^{L_{[a_2,b_2]}} = L_{[9,12]}^{L_{[6,11]}} = L_{[10.6875,12.5625]};$$

$$L_{[a_{res_3},b_{res_3}]} = L_{[a_{res_2},b_{res_2}]}^{L_{[a_3,b_3]}} = L_{[10.6875,12.5625]}^{L_{[0,10]}} \approx$$

$$\approx L_{[11.0769,12.6559]},$$

and thus:

$$L_{[a_{res},b_{res}]} = L_{[a_{res_3},b_{res_3}]} = L_{[11.0769,12.6559]}.$$

The resulting interval with the original ones is shown in Fig. 6.

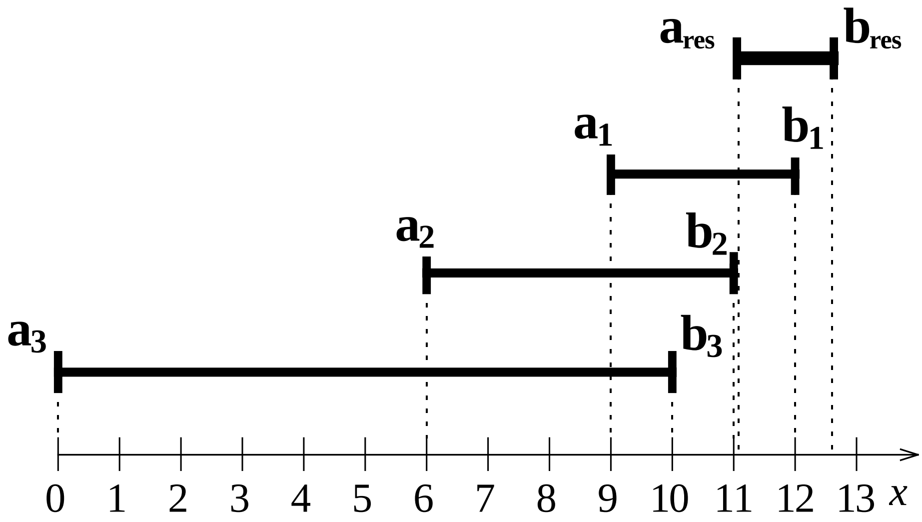

Fig. 6: The resulting interval and the original ones in the example

One can see that the resulting interval has no common points with two of the three original intervals. This happens because the decontextualization process takes into account the trend caused by noise in the estimation context.

3 A Trend Classification

In this chapter we consider one classification of trends into seven different groups of trends. This classification is used to group together intervals based on the relations of their endpoints and we prove that the previous step of decontextualization (Definition 2.4) gives as a result an interval that belongs to the same group as the intervals participating into the decontextualization process.

Definition 3.1:

There are seven groups of trends named as trends with *direction* dir_k and *power* pow_k (marked $L_k^{dir_k\ pow_k}$) as presented in Table 1. Each pair of interval estimations $L_{[a_i,b_i]}, L_{[a_j,b_j]} \in L_{[a_0,b_0]}, i \neq j$, belonging to the same group $L_k^{dir_k\ pow_k}$ keep the sign of $\Delta a + \Delta b$, Δa, and Δb where $\Delta a = a_j - a_i$, $\Delta b = b_j - b_i$.

The direction of a trend group is: *left ('l')*, *center ('c')*, or *right ('r')*, and it is defined by the sign of $\Delta a + \Delta b$:

$$(\Delta a + \Delta b > 0) \Rightarrow dir_k = 'l';$$

$$(\Delta a + \Delta b = 0) \Rightarrow dir_k = 'c';$$

$$(\Delta a + \Delta b < 0) \Rightarrow dir_k = 'r'.$$

The power of a trend group is: *slow ('<')*, *medium ('=')*, or *fast ('>')* and it is defined by the signs of Δa and Δb by the following way:

$$((\Delta a < 0) \textit{ and } (\Delta b > 0)) \Rightarrow pow_k = '<';$$

$$((\Delta a = 0) \textit{ or } (\Delta b = 0)) \Rightarrow pow_k = '=';$$

$$((\Delta a > 0) \textit{ or } (\Delta b < 0)) \Rightarrow pow_k = '>'.$$

In Fig. 7 there are three examples of trend groups with left direction and power: slow (a), medium (b), and fast (c).

Table 1: Trends of uncertainty

Trend	Direction→	left	central	right
Power↓	Restrictions	$\Delta a + \Delta b > 0$	$\Delta a + \Delta b = 0$	$\Delta a + \Delta b < 0$
slow	$(\Delta a < 0) and$ $(\Delta b > 0)$	$L^{l<}$	$L^{c<}$	$L^{r<}$
medium	$(\Delta a = 0) or$ $(\Delta b = 0)$	$L^{l=}$	does not exist	$L^{r=}$
fast	$(\Delta a > 0) or$ $(\Delta b < 0)$	$L^{l>}$	does not exist	$L^{r>}$

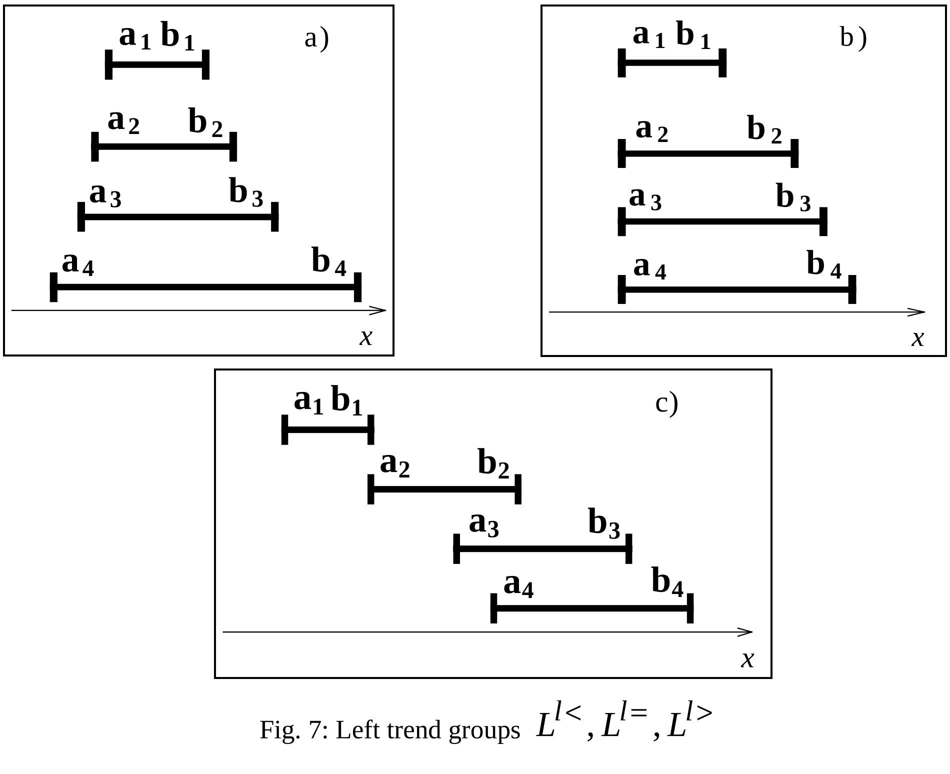

Fig. 7: Left trend groups $L^{l<}, L^{l=}, L^{l>}$

Theorem 3.1:

When $L_{[a_i,b_i]}^{L_{[a_j,b_j]}} = L_{[a_{res},\ b_{res}]}$, then the interval estimation $L_{[a_{res},b_{res}]}$ belongs to the same trend group as the interval estimation $L_{[a_i,b_i]}, L_{[a_j,b_j]}$.

Proof:

The interval estimations $L_{[a_i,b_i]}, L_{[a_j,b_j]}$ belongs to the group according to the signs:

$$sign(\Delta a + \Delta b) = sign(a_j - a_i + b_j - b_i),$$

$$sign(\Delta a) = sign(a_j - a_i), \text{ and}$$

$$sign(\Delta b) = sign(b_j - b_i).$$

To prove the theorem it is necessary to prove that:

$$sign(\Delta a' + \Delta b') = sign(a_i - a_{res} + b_i - b_{res}) = \\ = sign(\Delta a + \Delta b);$$

$$sign(\Delta a') = sign(a_i - a_{res}) = sign(\Delta a);$$

$$sign(\Delta b') = sign(b_i - b_{res}) = sign(\Delta b).$$

According the Definition 2.4:

$$a_{res} = a_i + \frac{u_i^2 \cdot (a_i - a_j)}{u_j^2 - u_i^2} = a_i - \frac{u_i^2}{u_j^2 - u_i^2} \cdot \Delta a;$$

$$b_{res} = b_i - \frac{u_i^2}{u_j^2 - u_i^2} \cdot \Delta b.$$

Thus:

$$\left.\begin{array}{l} (0 < u_i < u_j) \Rightarrow (\frac{u_i^2}{u_j^2 - u_i^2} > 0); \\ \Delta a' = a_i - a_{res} = \frac{u_i^2}{u_j^2 - u_i^2} \cdot \Delta a; \\ \Delta b' = b_i - b_{res} = \frac{u_i^2}{u_j^2 - u_i^2} \cdot \Delta b; \\ \Delta a' + \Delta b' = \frac{u_i^2}{u_j^2 - u_i^2} \cdot (\Delta a + \Delta b) \end{array}\right\} \Rightarrow$$

$$\Rightarrow sign(\Delta a') = sign(\Delta a);$$

$$sign(\Delta b') = sign(\Delta b);$$

$$sign(\Delta a' + \Delta b') = sign(\Delta a + \Delta b).$$

The Theorem 3.1 shows that the step of decontextualization gives resulting interval that belongs to the same group of trends as the original intervals.

4 Deriving a Resulting Interval from Several Trends

In a common case it is possible that several different trends can be derived from the same set of intervals. In this chapter we discuss one way of deriving resulting interval when there exist several trends among the original intervals.

Each pair of intervals can define a trend. We require that each pair participate once and only once in some of trends. This means that the number *m* of different trends cannot be more than the number of different interval pairs in the set of intervals $L = L_{[a_i,b_i]},\ i = 1,\ldots n$ given as opinions by knowledge sources:

$$m \le C_n^2.$$

Each trend can in general case include more than one interval and each interval can support several more than one trend.

Definition 4.1:

Let us suppose that the set *L* of interval opinions $L_{[a_i,b_i]},\ i = 1,\ldots n$ is divided into *m* trends $L_k,\ k = 1,\ldots m$.

The *support* S_k for the trend L_k is calculated as follows:

$$S_k = q_{res}^k \cdot \sum_{\forall i, L_{[a_i,b_i]} \in L_k} \frac{1}{N_i},$$

where q_{res}^k is the quality of the result $L_{[a_{res}^k, b_{res}^k]}$,

N_i is the number of different trends that includes the opinion $L_{[a_i,b_i]}$.

As one can see the Definition 4.1 gives more support for the trend that includes more intervals and the support of each interval is divided equally between all the trends that include this interval.

Definition 4.2:

Let the set of original interval estimations $L = L_{[a_i, b_i]}$, $i = 1, \ldots, n$ consists of m different trends L_k, $k = 1, \ldots, m$ with their resulting interval opinions $L_{[a^k_{res}, b^k_{res}]}$ and support S_k.

Then the resulting opinion $L_{[a^L_{res}, b^L_{res}]}$ for the original set of interval opinions is derived using following formulas:

$$a^L_{res} = \frac{a^1_{res} \cdot S_1 + a^2_{res} \cdot S_2 + \ldots + a^m_{res} \cdot S_m}{S_1 + S_2 + \ldots + S_m};$$

$$b^L_{res} = \frac{b^1_{res} \cdot S_1 + b^2_{res} \cdot S_2 + \ldots + b^m_{res} \cdot S_m}{S_1 + S_2 + \ldots + S_m}.$$

Thus the resulting interval is expected to be closer to the result of those trends that have more support among the original set of intervals.

5 Conclusion

This paper discusses one approach to handle interval uncertainty in estimation of some domain parameter. The case is considered when the estimation is made by multiple knowledge sources in a context of a trend caused by possible noise. The approach is based on an assumption that if a knowledge source guarantees less measurement error (estimation interval is shorter), then this source in the same time is more resistant against the effect of possible noise. In this paper we discussed one way to decontextualize knowledge given under misleading noise when this basic assumption holds. We defined different groups of trends among estimated intervals. We introduced one way how to take into account several trends that exist among the original intervals when one resulting interval is produced. One of the most important messages of this paper is as follows. If you have several opinions (estimations, recognition results, solutions etc.) with different values of uncertainty you can select and use the most precise one or take weighted mean value of these opinions. However it seems more reasonable to order opinions from the worst to the best one and try to recognize a trend of uncertainty, which (if exists) helps you to derive an opinion more precise than the best one. Further research is needed in different application areas to evaluate practical results of such assumption and algorithms of the decontextualization process.

Acknowledgement: This research is partly supported by the Academy of Finland.

References

[1] Alefeld G., Herzberger J., *Introduction to Interval Computation*, Academic Press, 1983.

[2] Benford S., Snowdon D., Greenhalgh C., Ingram R., Knox I., Brown C., VR-VIBE: A Virtual Environment for Co-operative Information Retrieval, *Eurographics'95*, Maastricht, Netherlands, 1995, pp. 349-360.

[3] Brezillon P., Context Needs in Cooperative Building of Explanations, In: *Proceedings of the First European Conference on Cognitive Science in Industry*, 1994, pp. 443-450.

[4] Brezillon P., Abu-Hakima S., Using Knowledge in its Context: Report on the IJCAI-93 Workshop, *AI Magazine*, Vol. 16, 1995, pp. 87-91.

[5] Brezillon P., Cases E., Cooperating for Assisting Intelligently Operators, In: *Proceedings of Actes International Workshop on the Design of Cooperative Systems*, 1995, pp. 370-384.

[6] Brooks T., Topical Subject Expertise and the Semantic Distance Model of Relevance Assessment, *Journal of Documentation*, Vol. 51, No. 4, 1995, pp. 370-387.

[7] Buvac S., Buvac V., Mason I., Metamathematics of Contexts, *Fundamenta Informaticae*, Vol. 23, 1995, pp. 263-301.

[8] Chalmers M., Chitson P., Bead: Explorations in Information Visualization, In: *Proceedings of SIGIR'92*, ACM Press, 1992, pp. 330-337.

[9] Cugini J., Piatko C., Laskowski S., Interactive 3D Visualization for Document Retrieval, *Contribution of the National Institute of Standards and Technology*, 1998, Available in WWW: http://zing.ncsl.nist.gov/~cugini/uicd/viz.html.

[10] Foo N., Garner B., Rao A., Tsui E., Semantic Distance in Conceptual Graphs, In: L. Gerhotz (ed.), *Current Directions in Conceptual Structure Research*, Ellis Horwood, 1992, pp. 149-154.

[11] Gerasin S., Kaikova H., Terziyan V., An Interval-Based Approach to Multiple Experts Opinions Processing: Successive Severance Method, *Bionics Problems Journal*, State University Publishers, 1990, Vol. 44, pp. 41-46 (in Russian).

[12] Godo L., Mantaras R., Sierra C., Verdaguer A., MILORD: The Architecture and the Management of Linguistically Expressed Uncertainty, *International Journal of Intelligent Systems*, Vol. 4, 1989, pp. 471-501.

[13] Guha R., *Contexts: A Formalization and Some Applications*, Stanford Ph.D. Thesis, 1991.

[14] Heylighen F., Bootstrapping Knowledge Representation: from Entailment Meshes via Semantic Nets to Learning Webs, *International Journal of Human-Computer Studies* (submitted), 1997, Available in WWW: http://pespmc1.vub.ac.be/Papers/BootstrappingPask.html.

[15] Jorgensen P., *P Jorgensen's COM515 Project Page*, 1996, Available in WWW: http://jorg2.cit.bufallo.edu/COM515/project.html.

[16] Kreinovich V., Data Processing Beyond Traditional Statistics: Applications of Interval Computations, In: *Proceedings of International Workshop on Applications of Interval Computations - APIC'95*, El Paso, 1995, pp. 13-21.

[17] Krohn U., VINETA: Navigation Through Virtual Information Spaces, In: T. Cartaci (ed.), *AVI: Advanced Visual Interfaces*, Gubbio, Italy, ACM, 1996.

[18] Kulisch U., Miranker W., *Computer Arithmetic in Theory and Practice*, Academic Press, 1981.

[19] Lynch K., Chen H., Knowledge Discovery from Historical Data: An Algorithmic Approach, *In: Proceedings of the 25th Annual Hawaii International Conference on System Sciences*, Kaui, 1992, pp. 70-79.

[20] McCarthy J., Notes on Formalizing Context, In: *Proceedings of 13-th International Joint Conference on Artificial Intelligence*, 1993, pp. 555-560.

[21] Moore R., *Interval Analysis*, Prentice Hall, Englewood Cliffs, NJ, 1966.

[22] Moore R., *Methods and Applications of Interval Analysis*, SIAM, Philadelphia, Pennsylvania, 1979.

[23] Nguyen H., Walker E., Interval Fuzzy Data Processing: Case when Degree of Belief is Described by an Interval, In: *Proceedings of International Workshop on Applications of Interval Computations - APIC'95*, El Paso, 1995, p. 158.

[24] Olsen K., Korfhage R., Sochats K., Spring M., Williams J., Visualization of a Document Collection: The VIBE System, Information Processing and Management, *Pergamon Press Ltd.*, Vol. 29, No. 1, 1993, pp. 69-81.

[25] Rada R., Mili H., Bicknell E., Blettner M., Development and application of a metric on semantic nets, *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 19, No. 1, 1989, pp. 17-30.

[26] Rocha L., Luis M., Fuzzification of Conversation Theory, In: F. Heylighen (ed.), *Principia Cybernetica Conference*, Free University of Brussels, 1991.

[27] Tailor C., Tudhope D., Semantic Closeness and Classification Schema Based Hypermedia Access, In: *Proceedings of the 3-rd International Conference on Electronic Library and Visual Information Research (ELVIRA'96)*, Milton, Keynes, 1996.

[28] Taylor W., Weimann D., Martin P., Knowledge Acquisition and Synthesis in a Multiple Source Multiple Domain Process Context, *Expert Systems with Applications*, Vol. 8, 1995, pp. 295-302.

[29] Wilson D., Martinez T., Improved Heterogeneous Distance Functions, *Journal of Artificial Intelligence Research*, Vol. 6, 1997, pp. 1-34.

Vagan Terziyan 1958, received his Engineer-mathematician degree on Applied Mathematics in 1981 from Kharkov State Technical University of Radioelectronics (KhTURE). He became Candidate of Technical Sciences in 1985 on Technical Cybernetics and Information Theory and Doctor of Technical Sciences in 1993 on Information Systems and Artificial intelligence. He is acting as Professor since 1994 and as the Head of the Department of Artificial Intelligence and Information Systems since1997 in KhTURE. Subject area - Artificial Intelligence: Knowledge Representation and Reasoning, Context in AI, Data Mining and Knowledge Discovery, Machine Learning, Expert Systems, Cooperative Decision-Making, Medical Informatics. He is the Scientific Head of the Metaintelligence Laboratory in KhTURE, the Head of Cooperation Program with University of Jyvaskyla (Finland), the Head of the Research Department of the Laboratory of Distance Learning and Virtual Education (Ministry of Education of Ukraine), member of three scientific boards. He made his visiting research and professorship in the University of Jyvaskyla in 1996 (March-December) due to the Grant by the Center for International Mobility, in 1997 (August-December) and 1998 (August - December) due to the Grants by the Academy of Finland. He has near 100 scientific publications, more than 30 in internationally recognized magazines and conferences.

Seppo Puuronen 1949, received his Ph.D. on Computer Science in 1988 from the University of Jyvaskyla (Finland). He is acting as assistant professor of Computer Science at the University of Jyvaskyla and he acted as Professor, Associate Professor and Lecturer of Computer Science at Jyvaskyla and Laapenranta University of Technology. He acted as Head of the Department of Computer Science and Information Systems at the University of Jyvaskyla 1988-1989, 1990-1991 and 1997. He was the member of Program Committee of the Finnish AI Symposium in 1988, as a tutorial chair of CAISE'95, and as a chair of the Finnish AI Symposium 1998. He became Honor-Doctor of the Kharkov State Technical University of Radioelectronics in 1997. His main research interests are Knowledge Engineering, Expert Systems, and Computers in Education. He has published 40 scientific papers, 30 in internationally recognized magazines and conferences.

Helen Kaykova 1960, received her Engineer-system-analytic degree on Automatic Control Systems in 1982 from Kharkov State Technical University of Radioelectronics (KhTURE). She became Candidate of Technical Sciences in 1989 on Technical Cybernetics and Information Theory. She is acting as Docent of Software Department since 1992 and of the Department of Artificial Intelligence and Information Systems since 1997 and as the Vice-Rector on International Research Cooperation since 1996 in KhTURE. Subject area - Artificial Intelligence: Logic, Temporal Reasoning, AI and Statistics, Cooperative Decision-Making. She is the Director of the Metaintelligence Laboratory in KhTURE, the Coordinator of Cooperation Program with University of Jyvaskyla (Finland), the Head of the International Department of the Laboratory of Distance Learning and Virtual Education (Ministry of Education of Ukraine). She made her visiting research in the University of Jyvaskyla in 1996 (May-December) due to the Grant by the Ministry of Education of Finland, in 1997 (August-December) and 1998 (August - December) due to the Grants by the Academy of Finland. She has near 40 scientific publications, 14 in internationally recognized magazines and conferences.