How Similar is VERY YOUNG to 43 Years of Age?

On the Representation and Comparison of Polymorphic Properties

Werner Dubitzky, Alfons Schuster, John G. Hughes, David A. Bell*, Kenneth Adamson§
University of Ulster, Jordanstown
Northern Ireland Bio-Engineering Centre*
Faculty of Informatics
Co. Antrim BT37 OQB Northern Ireland

Abstract

Intelligent computer systems rely on more or less complex computational entities that *represent* occurrences and events in the real world. Usually, such entities are formed from representational primitives called properties, attributes, features, etc. To reflect varying degrees of uncertainty, originating from human judgement and the intrinsic nature of the world, such property values occur as more or less vague linguistic symbols or exact numeric expressions. Determining similarity between two properties is usually done on either the symbolic or the numeric level. This seems to be too restrictive for case-based reasoning and similar approaches as these often face mixed specifications. In this paper we propose a flexible and systematic scheme for representing crisp properties and two types of fuzzy properties. It also provides a consistent mechanism to establish similarity scores for the various instance combinations.

1. Introduction

The main thread of this paper is that of *conceptual similarity* (or distance). Two quotes by I. Kant and W.V. Quine succinctly convey the essence and significance of this idea. 'A *concept without a percept is empty; a percept without a concept is blind.* and *'There is nothing more basic to thought and language than our sense of similarity; our sorting of things into kinds.'*

Many knowledge and information representation models employ the notion of a *property* by some means or other. For example, object-oriented database systems *(attribute* or *atomic object),* multiple criteria decision making *(criterion),* relational databases *(field* or *attribute),* statistics *(statistical variable),* frame-based systems *(slot),* case-based reasoning *(feature),* concept theory *(property),* and so on. Properties are usually viewed as representational primitives that provide a 'vehicle' to form more complex and more abstract structures like *classes, complex objects, alternatives, relations, records, tables, frames, cases, concepts, concept exemplars,* and so forth. It is these higher level entities that are used in computer systems to represent and capture the essence of real-world occurrences and events. Frequently, the underlying processes of knowledge-based and other systems require the (conceptual) comparison of such computational units. Comparison methods of this kind establish similarity scores based on the similarity of such atomic properties [Chen and Hwang, 1992].

Intelligent computer systems involve application-geared people at both the knowledge acquisition stage as well as at the problem-solving and decision-making stage. The reasoning quality of such systems hinges upon the adequacy and appropriateness with which experts and users define the characteristics of properties and specify property instances. Inevitably, because of human involvement and the intrinsic nature of real-world circumstances, uncertainty is manifest in the definitions and specifications of many properties [Zadeh, 1973; Klir *et al,* 1988].

This work proposes the concept of a *polymorphic property.* It enables the specifier of instances of the *same* property type to choose from three value representations according to prevailing uncertainty and his or her level of expertise. Further, to enable proximity-based reasoning models to take advantage of this representation device, a consistent method is presented that allows the computation of similarity between the various value combinations. This method has been developed with four objectives in mind: it should be intuitive, intellectually satisfying, easy to use, and computationally efficient.

The approach has been applied within a case-based reasoning (CBR) framework in the Coronary Heart Disease risk assessment domain. The major results of this study were that 1) Unassisted subjects (users) are able to specify their own case data in the presence of varying degrees of uncertainty. 2) the system can consistently process the typical CBR inference steps on cases whose property values were provided in crisp numeric and vague linguistic values. And 3) A benchmark test with 10,000 cases shows that the method is highly efficient.

It is claimed that the method's mechanisms are applicable to a wider range of systems and applications, especially in practical knowledge engineering environments.

The paper is organised as follows: Section 2 briefly discusses some related approaches and illustrates the need for more flexible uncertainty handling. Section 2.1 introduces and defines the notion of polymorphic property values. Sections 3.1 to 3.3 develop the various comparison methods. Section 4 briefly mentions the CBR application in which this approach has been tested, and a performance experiment that has been carried out. Section 5 ends with conclusions.

2. Multiple-Format Property Values

In CBR systems uncertainty is pervasive in, amongst other things, the features used to describe the cases [Dutta *et al.*% 1991]. In such and other systems uncertainty may originate from various sources [Zadeh, 1973; Chen and Hwang, 1992; Baldwin *et al.,* 1995]:

- unquantifiable information (judgement of individual): e.g., a patient's *age* can be easily determined while the *criticality* of his or her condition is not readily quantifiable,
- incomplete information (inexact measurement): e.g., a patient's *blood pressure* may be measured as 'about 145 mmHg', but not as 'exactly 145 mmHg',
- non-obtainable information (data too expensive to obtain): sometimes crisp data is in principle obtainable, but the cost is too high, e.g., a patient's *family history of high cholesterol.* It may, however, be possible to get a useful approximation of that data,
- partial ignorance (partially known facts about a phenomenon): some facts may only be partially known, e.g., the (average) *number of cigarettes per day* a patient smoked over the last 20 years.

Various data/knowledge models handle some of these forms of uncertainty in some way or another. For example, multiple attribute decision making systems [Chen and Hwang, 1992] (fuzzy sets), relational data models [Buckles *et* a/., 1982; Petry, 1996] (fuzzy sets, rough sets), object-oriented information models [George *et al.,* 1993] (fuzzy sets), and [Baldwin *et al,* 1995] (fuzzy sets, Evidence Theory). And in CBR [Tirri *et al.,* 1996] (Bayesian), [Dubitzky *et* a/., 1996a/b] (fuzzy sets, Evidence Theory), and [Dutta *et al.,* 1991] (fuzzy sets).

Common to these approaches is that a particular property may only take values of a single type. Usually these values are expressed by *linguistic symbols* (associated with a fuzzy set or a similarity matrix), *or* [exclusive or!] *crisp numbers.* This restriction, however, does not reflect the practices in many real-world problem-solving and decision-making environments, where data are provided (mainly by people) in whatever format is appropriate or available at the time. This may result in values of a particular property type occurring in both fuzzy and crisp formats.

Consider, for example, a printer product support help desk scenario where customers are asked how many times a certain failure had occurred. Some may have actually counted the breakdowns, others may only be able to give an approximate estimate. Or imagine a healthcare system asking advice-seeking users to state their blood cholesterol concentrations. Some users may have recently had their cholesterol checked, and remember exact readings (or indeed a value like *low* reported to them by their medic). Others may only be able to recall, vaguely or accurately, measurements given to them in the past. Yet another group may simply come up with a more or less informed guess which is likely to be expressed linguistically rather than numerically.

To facilitate properties that consistently represent and process (similarity assessment) such polymorphic instance values, the notion of a property *concept frame* is introduced.

2.1 Property Concept Frame

In order to equip a property (type) with the capability to have instances of different form, the property (type), and its instances, are associated with a *concept frame.* The concept frame provides the representation platform to model the relationships between the various value formats, thus enabling the computation of cross-format similarity scores.

To be more precise, a concept frame serves as a unifying representation formalism for three property value formats: *real number, linguistic term,* and *fuzzy predicate.* These provide a means to express property instances at various levels of certainty and expertise. The concept frame of the property *cholesterol* (abbreviated by *chol)* depicted in Figure 1 should help to illustrate this idea.

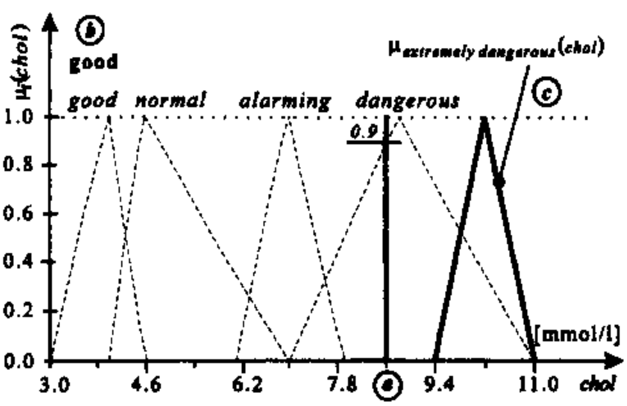

Figure 1: Property concept frame and three actual property values.

Basically, a concept frame consists of a finite *universe of discourse* (in Figure 1, the blood cholesterol concentration range 3.0... 11.0 mmol/1), and a set of pre-defined concepts. These concepts are represented via *linguistic symbols* and the corresponding membership functions which essentially define *fuzzy sets.* Membership functions are referred to as *v-functions;* v stands for value.

So the concepts in the *cholesterol* concept frame example (see diagram) are defined by the linguistic terms *good*, *normal*, *alarming*, and *dangerous*, and the v-functions (dashed triangles in Figure 1) $\mu_{good}(chol)$, $\mu_{normal}(chol)$, $\mu_{alarming}(chol)$, and $\mu_{dangerous}(chol)$.

By means of the linguistically expressed concepts, whose semantics are reflected via the location and shape of the corresponding v-functions, a domain expert effectively conceptualises the universe of discourse according to his or her experience. Thus, the universe of discourse together with the linguistic terms and v-functions form the conceptual frame of discernment, or concept frame, within which *all* instances of the associated property are interpreted.

Figure 1 also demonstrates three examples of actual cholesterol *values* (objects depicted in bold style) indicated by a, b, and c. The value a is given in real number format, $a := \langle 8.6 \rangle$ [mmol/l], b in linguistic term format, $b := \langle good \rangle$, and c in fuzzy predicate format, $c := \langle extremely\ dangerous, \mu_{extremely\ dangerous}(chol) \rangle$.

The real number format is employed where quantifiable data is readily available. However, as it will be seen later, real numbers are reflected upon the concepts (v-functions) mapped out on a concept frame. For example, the crisp value $a := \langle 8.6 \rangle$ maps onto the *dangerous* concept as follows: $\mu_{dangerous}(a) = 0.90$. Note, on the conceptual plane the value a is perceived as $\langle dangerous, 0.90 \rangle$. Real number format values must be drawn from the concept frame's universe of discourse.

To specify a property value linguistically, the specifier of the value may chose any of the linguistic terms associated with the v-functions of a particular concept frame. For example, *(dangerous).* At first glance this may appear rather restrictive. However, the concepts defined on a concept frame are carefully established according to an expert's decision-making and problem-solving experience with the property and application in question. This ensures some degree of consistency, especially when inexperienced users specify values. An additional format (see below) provides more freedom and flexibility in cases where crisp data is not available and the set of pre-defined linguistic symbols does not seem to include an appropriate symbol.

Finally, the fuzzy predicate format. It provides a flexible mechanism to freely specify a value by introducing and defining a *new* concept and its semantics. He or she does so by stating a linguistic expression, naming the actual concept, and defining the corresponding membership function which characterises the meaning of the concept. An example for a fuzzy predicate value is given by the value *c* depicted in Figure l. The intention is to equip the more experienced user (including the domain experts themselves) with a flexible and powerful representation 'vehicle' to express values more appropriately in accordance with the prevailing uncertainty.

Definition 1 A concept frame associated with a polymorphic property defines a *universe of discourse U* and a *concept domain C. C* is composed of linguistic symbol/v-function pairs as follows:

$$C \underset{def}{=} \{ (A_1, \mu_{A_1}(x)), (A_2, \mu_{A_2}(x)), \ldots, (A_n, \mu_{A_n}(x)) \}$$

where A_i denotes a linguistic label, and each $\mu_{A_i}(x)$ is a membership function or v-function defining a fuzzy set over *U,* such that $\mu_{A_i}(x) : U \rightarrow [0,1]$, for $x \in U \subseteq$ R, and $1 \leq n \leq 9$. (note, all v-functions must define *normalised* and *convex* fuzzy sets)

Definition 2 The *value* formats by means of which instances of a polymorphic property may be specified comprise the real number (/WV), the linguistic term *(LT),* and the fuzzy predicate format *(FP).* These are succinctly defined as follows:

$$RN \underset{def}{=} \langle x \rangle, LT \underset{def}{=} \langle A \rangle, FP \underset{def}{=} \langle B, \mu_B(x) \rangle$$

where *A* denotes a linguistic symbol that must be drawn from the terms defining the concept domain, i.e., $A \in \{A_1, A_2, \ldots, A_n\}$, *B* denotes a linguistic term naming a concept whose semantics is defined by the membership function $\mu_B(x)$, such that $\mu_B(x)$: $U \rightarrow [0,1]$, and $x \in U$, $B \notin \{A_1, A_2, \ldots, A_n\}$.

3. Comparing Polymorphic Property Values

Many decision-making and problem-solving models require the (conceptual) comparison of more or less complex entities involved in the process [Chen and Hwang, 1992; Dubitzky *et al.,* 1996b]. Eventually, such algorithms come down to establishing similarity or distance between the constituent properties used to describe the entities in question. So providing a system or knowledge engineer with a powerful representation mechanism, such as the polymorphic property discussed above, is not enough. One must also put at his or her disposal a scheme that allows the systematic comparison of instances of such properties.

Because of symmetry, i.e., *sim(x,y) = sim(y,x)),* a total of six possible value format combinations need to be considered, namely (LT,LT), *(FP,FP), (LT,FP),* (RN,RN). *(RN,LT),* and *(RN,FP).* (where *RN* = real number, *LT* = linguistic term, and *FP* = fuzzy predicate format) These can be grouped into 1) fuzzy/fuzzy format comparison: *(LT,LT), (FP,FP),* and (LT,FP), 2) crisp/crisp: *(RN,RN),* and 3) crisp/fuzzy: *(RN,LT)* and *(RN,FP).* As the method for each group is in principle the same, only one combination per group has to be investigated.

3.1 Comparing Fuzzy Properties

Here, a fuzzily formatted property value is represented via a linguistic term and a associated membership function (essentially defining a fuzzy set). This is the case for both the linguistic term format as well as the fuzzy predicate format. Since fuzzy sets, or for that matter, fuzzy numbers represent many possible real numbers (with different membership degrees) they do not always yield a totally ordered set. This makes comparison a non-trivial affair. Various approaches have been proposed in the literature [Chen and Hwang, 1992]. Because of its simplicity, intuitive appeal, and effectiveness, Chen and Hwang's *crisp score method* [Chen and Hwang, 1992, 465-486] is adopted here to defuzzify fuzzy sets.

Given the universe of discourse U, a fuzzy set A is characterised by a membership function $\mu_A(x)$, such that $\mu_A(x) : U \rightarrow [0,1]$, where $x \in U$. Then, depending on shape/location of $\mu_A(x)$ in U, the crisp score method derives a crisp score $s(A)$ from A, such that $s(A) \in [0,1]$. And in case of fuzzy numbers $U \subseteq \mathbf{R}$.

Now, based on crisp scores, the idea of a *basic concept distance* is defined. It will play a part in the computation of concept distances for most of the possible value combinations.

Definition 3 Let A and B denote linguistic terms naming the corresponding concepts (linguistic term, fuzzy predicate values) on a concept frame, then the *basic concept distance* $\delta(A,B)$ between A and B is defined by the crisp scores $s(A)$ and $s(B)$ as follows:

$$\delta(A,B) = \left| s(A) - s(B) \right| \qquad (1)$$

such that A, $B \in \{A_1, A_2, \ldots, A_n\}$ for linguistic term values, and A, $B \in$ {newly introduced labels} for fuzzy predicate values.

For example, the basic concept distance between the linguistic *cholesterol* values *good* and *alarming* yields $\delta(good,alarming) = | s(good) - s(alarming) | = 0.56$.

As the (general) concept distance $d(A,B)$ between two fuzzy value format values (linguistic term, fuzzy predicate) is identical to the corresponding basic concept distance $\delta(A,B)$, i.e., $d(A,B) = \delta(A,B)$, it is not defined separately.

3.2 Comparing Real Number Properties

The real world can be thought of as consisting of objects and events which are characterised by continuous numeric values. People, on the other hand, represent and process their knowledge by means of symbols. In order to make it possible to compare real number format values with any of the other three formats they are 'mapped' onto the concept arrangement (essentially *expressed* via symbols) of the corresponding concept frame. The vehicle to relate the symbolic and numeric levels is provided by the fuzzy sets (v-functions) used to represent the meaning of the linguistically described concepts. This means, for example, that the value a := (8.6) becomes $\langle$*dangerous*/0.90$\rangle$.

Essentially, the computation of a *conceptual* distance measure $d(a,b)$ between two real number values a and b is determined by the basic concept distances of the concepts on the concept frame that are 'affected' by a and b, *and* the degree to which these concepts are reflected.

First, the situation is considered where the two values a and b map onto a single concept A. The initial observation is that, independent of a and b, the basic concept distance $\delta(A,A)$ is zero; see Definition 3. Clearly then, with respect to concept A, the conceptual difference between a and b is measured on how a and b differ with respect to their conformance with concept A. This, of course, is computed by $| \mu_A(a) - \mu_A(b) |$. A crucial consequence of this is that, conceptually, two crisp numeric values a and b may be identical, i.e., concept distance $d(a,b) = 0$, if $\mu_A(a) = \mu_A(b)$ holds, in spite of $a \neq b$.

Second, what happens if the value a maps onto a concept A, and b onto a *different* concept B (i.e., $a \neq b$ and $A \neq B$)? Here, according to Definition 3, the basic concept distance $\delta(A,B)$ is always greater than zero. Somehow, one would think, the concept distance $d(a,b)$ between a and b should depend on $\delta(A,B)$. Now let a and b take on values such that $\mu_A(a) = 1$ and $\mu_B(b) = 1$, indicating full compliance of a and b with the concepts A and B respectively. In this special case, the concept distance $d(a,b)$ should be equal to the basic concept distance $\delta(A,B)$, that is, $d(a,b) = \delta(A,B)$. This state of affairs is brought about by treating $\langle A/1.00 \rangle$ as $\langle A \rangle$, and $\langle B/1.00 \rangle$ as $\langle B \rangle$. To put it another way: $(\mu_A(x_1) = 1) \rightarrow A$, and $(\mu_B(x_2) = 1) \rightarrow B$. This effectively increases the degree of uncertainty in the system, for the sake of being able to compare crisp numeric with vague linguistic values.

Preserving these two principles, the comparison method is based on two functions which are derived from the v-functions describing the concepts on a concept frame. These functions are $\delta^+(x)$ (solid graph in Figure 2b) and $\delta^-(x)$ (dashed graph in Figure 2b).

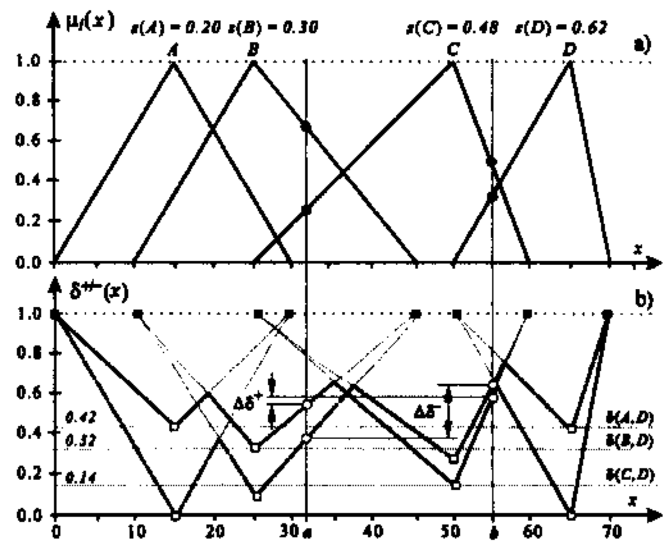

Figure 2: Concept frame and $\delta^+(x)$ and $\delta^-(x)$.

Function $\delta^+(x)$ describes the *directed conceptual distance* measured from the concept with the lowest crisp score (A) to the one with the highest (D), and $\delta^-(x)$ from that with the highest crisp score (D) to the one with the lowest (A). The functions are constructed from the (directed) basic concept distances $\delta(A,D)$, $\delta(B,D)$, $\delta(C,D)$, $\delta(D,D)$, for $\delta^+(x)$, and $\delta(D,A)$, $\delta(C,A)$, $\delta(B,A)$, $\delta(A,A)$, for $\delta^-(x)$ (white squared dots). For example,

$\delta(A,D) = 0.42$, $\delta(B,D) = 0.32$, $\delta(C,D) = 0.14$, $\delta(D,D) = 0$. The concept distance between two values a and b is then computed via the *differentials* $\Delta\delta^{+}(a,b)$ and $\Delta\delta^{-}(a,b)$ as: $d(a,b) = (\Delta\delta^{+}(a,b) + \Delta\delta^{-}(a,b))/2$, where $\Delta\delta^{+}(a,b) = |\delta^{+}(a) - \delta^{+}(b)|$, and $\Delta\delta^{-}(a,b) = |\delta^{-}(a) - \delta^{-}(b)|$.

Definition 4 Let a and b denote real number format values in U, then the concept distance $d(a,b)$ is determined by the directed conceptual distances $\delta^{+}(x)$ and $\delta^{-}(x)$ as follows (see also Definition 5):

$$d(a,b) = \tfrac{1}{2}\{\,|\delta^{+}(a) - \delta^{+}(b)| + |\delta^{-}(a) - \delta^{-}(b)|\,\} \quad (2)$$

With the example values $a = 32$ and $b = 55$, depicted in Figure 2, the concept distance works out to $d(32,55) = 0.12$.

Figure 3 below illustrates the behaviour of the concept distance between two crisp numeric values. There, crisp *age* values from 0 to 100 are compared with the value 65 years. To demonstrate the relationship with the *age* concept frame, the v-functions (dashed graphs) and their corresponding linguistic terms are indicated in the diagram.

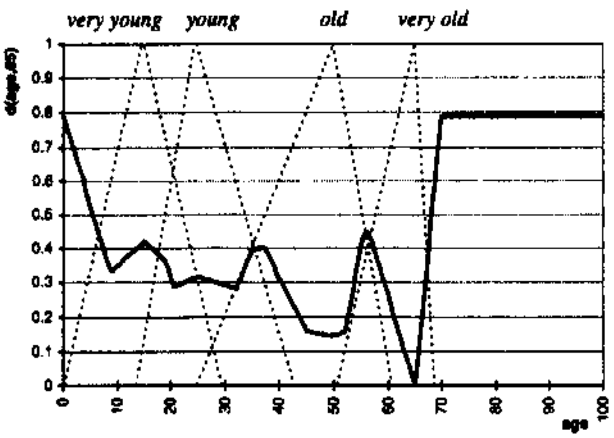

Figure 3: Concept distance between *age* and 65 years.

Definition 5 The functions $\delta^{+}(x)$ and $\delta^{-}(x)$ are defined by the functions $\rho_{A_1}(x), \rho_{A_2}(x), \ldots, \rho_{A_i}(x), \ldots, \rho_{A_n}(x)$, and $\lambda_{A_1}(x), \lambda_{A_2}(x), \ldots, \lambda_{A_i}(x), \ldots, \lambda_{A_n}(x)$ as follows:

$$\delta^{+}(x) = min[\,\rho_{A_1}(x), \rho_{A_2}(x), \ldots, \rho_{A_i}(x), \ldots, \rho_{A_n}(x)\,] \quad (3a)$$

$$\delta^{-}(x) = min[\,\lambda_{A_1}(x), \lambda_{A_2}(x), \ldots, \lambda_{A_i}(x), \ldots, \lambda_{A_n}(x)\,] \quad (3b)$$

where each $\rho_{A_i}(x)$ and $\lambda_{A_i}(x)$ relates to and is derived from the corresponding v-function $\mu_{A_i}(x)$ on the concept frame (A_1 is the concept with the lowest crisp score, and A_n the highest). In Figure 2b, $\rho_{A_i}(x)$ and $\lambda_{A_i}(x)$ are indicated by the lines connecting a white squared dot with two black squared dots.

The functions $\rho_{A_i}(x)$ and $\lambda_{A_i}(x)$ are obtained from the corresponding v-functions $\mu_{A_i}(x)$ by the linear mappings r and l as follows:

$$r : \mu_{A_i}(x) \rightarrow \rho_{A_i}(x) \quad (4a)$$

$$l : \mu_{A_i}(x) \rightarrow \lambda_{A_i}(x) \quad (4b)$$

such that $(x_L,\mu_{A_i}(x_L)=0) \rightarrow (x_L,\rho_{A_i}(x_L)=1)$, $(\hat{x},\mu_{A_i}(\hat{x})=1) \rightarrow (\hat{x},\rho_{A_i}(\hat{x})=\delta(A_i,A_n))$, $(x_R,\mu_{A_i}(x_R)=0) \rightarrow (x_R,\rho_{A_i}(x_R)=1)$, and $(x_L,\mu_{A_i}(x_L)=0) \rightarrow (x_L,\lambda_{A_i}(x_L)=1)$, $(\hat{x},\mu_{A_i}(\hat{x})=1) \rightarrow (\hat{x},\lambda_{A_i}(\hat{x})=\delta(A_i,A_1))$, $(x_R,\mu_{A_i}(x_R)=0) \rightarrow (x_R,\lambda_{A_i}(x_R)=1)$, for $x_L, x_R, \hat{x} \in U$. Where x_L and x_R refer to the left and right zero point and $\hat{x}$ to the height point of $\mu_{A_i}(x)$.

For example, from $\mu_C(x)$, third v-function of the concept frame shown in Figure 2a, the functions $\rho_C(x)$ and $\lambda_C(x)$ are obtained as:

$$\rho_C(x) = \begin{cases} \frac{1-\delta(C,D)}{(x_L=25)-(\hat{x}=50)}(x - x_L) + 1, \mathit{for}\ x_L \le x \le \hat{x} \\ \frac{1-\delta(C,D)}{(x_R=60)-(\hat{x}=50)}(x - x_R) + 1, \mathit{for}\ \hat{x} \le x \le x_R \end{cases}$$

$$\lambda_C(x) = \begin{cases} \frac{1-\delta(C,A)}{(x_L=25)-(\hat{x}=50)}(x - x_L) + 1, \mathit{for}\ x_L \le x \le \hat{x} \\ \frac{1-\delta(C,A)}{(x_R=60)-(\hat{x}=50)}(x - x_R) + 1, \mathit{for}\ \hat{x} \le x \le x_R \end{cases}$$

3.3 Comparing Crisp with Fuzzy Properties

Essentially, fuzzy predicate and linguistic term format values are represented, albeit not expressed, in the same way, namely via a fuzzy set (v-function, 'free' membership function). It is therefore sufficient to focus on the comparison between the crisp value format and linguistic term format values.

A crucial property of such a comparison must be that it never yields zero as a result. For d(*very young*,43) = 0 would imply identity and therefore suggest that *very young* = 43! This, however, would assert certainty where there is in fact no certainty, and would therefore render the scheme rather flawed.

The graph in Figure 4 illustrates the behaviour of the method by comparing the crisp *age* values 0 to 100 years with the vaguely specified value *very young*. It is interesting to note that the concept distances d(*age*,*very young*) are minimal around the *very young* region, but never reach zero! By way of digression, the white dot in the diagram indicates the answer to the question posed in the title of the paper, since *similarity*(*43*,*very young*) = 1 − d(*43*,*very young*) ≈ 0.77.

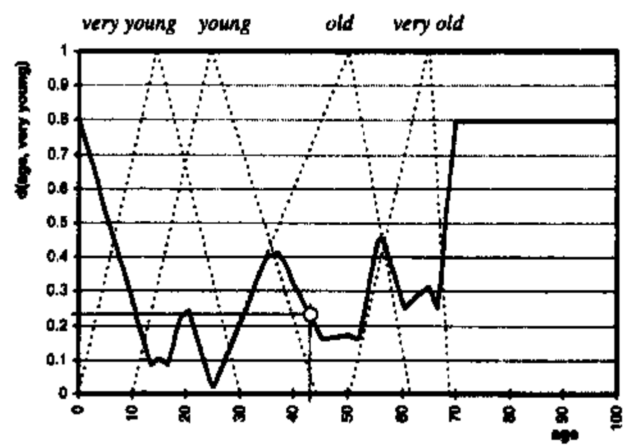

Figure 4: Concept distance between crisp *ages* and *very young*.

Defining the concept distance for this fuzzy-against-crisp case, three requirements are considered: 1) there should be no identity such that $d(fuzzy,crisp) = 0.0$, 2) the general tendency should be preserved; this means that the distance should decrease, if non-monotonously, for values closer to the *support* of fhe fuzzy set representing the fuzzy value, and 3) the method, obviously, should be sensitive to the shape/location of the membership function associated with the vague value.

Criterion 1) is achieved by deriving the distance form the sum of the (directed) basic concept distances $\delta(A_i,A_n)$ and $\delta(A_i,A_1)$ of the involved fuzzy value A_i,. (see also (4a) and (4b)) Without formal proof it is evident that $\delta(A_i,A_n) + \delta(A_i,A_1) > 0$ always holds (see Figure 2b).

Criterion 2) is inherently met by the definition of $\delta(A_i,A_n)$ and $\delta(A_i,A_1)$ which already reflect tendency or directedness.

And to satisfy criterion 3), the crisp score *s(Ai)*—which captures the membership function's shape and location on the concept frame—of the fuzzy value A_i is explicitly used.

Definition 6 Let A_i be a fuzzy property value (linguistic term or fuzzy predicate), and *x* a real number value. Then the concept distance *d(x,Ad* between A_i, and *x* is computed by (see also (2), (3a), and (3b)):

$$d(x,A_i) = \frac{1}{2}\left[\left|\delta^+(x) - \frac{s(A_i)+\delta(A_i,A_n)}{2}\right| + \left|\delta^-(x) - \frac{s(A_i)+\delta(A_i,A_1)}{2}\right|\right] \quad (5)$$

where $s(A_i)$ denotes A_i's crisp score, $\delta(A_i,A_n)$ and $\delta(A_i,A_1)$ denote the (directed) basic concept distances with respect to A_i, and $\delta^+(x)$ and $\delta^-(x)$ denote the directed conceptual distances, such that $x \in U$, for example:

$$d(43, very\ young) = \frac{1}{2}\left[\left|0.38 - \frac{0.20+0.42}{2}\right| + \left|0.48 - \frac{0.20+0}{2}\right|\right] = 0.23$$

4. A CBR Application and Some Benchmarks

The method has been tested in an experimental CBR system to give initial advice to subjects on their coronary heart disease (CHD) risk. The main point of this study was to have the subjects use the system and provide CHD-relevant data *without* first consulting a medic to establish various parameters such as *blood pressure, cholesterol,* or *anxiety.* To establish some idea about the method's performance 10,000 cases described by 21 polymorphic properties (7 concept frames with 5 v-functions, 7 with 7, and 7 with 9) have been seeded randomly with values of the three formats. On 100 different value distributions over all cases, the average time to compare all 10,000 cases was 11.74 seconds (minimum 7.22 sec, maximum 16.63 sec). The test configuration was a 90 MHz Pentium PC, 32 MB RAM on a Windows NT4 platform.

5. Conclusions

The polymorphic property approach presented in this paper led to encouraging results. It seems pertinent to applications (e.g., tele-medicine or help desk systems) involving a great deal of data that is provided in crisp as well as in fuzzy format. Further, performance tests indicate that the method does not incur exceptional overheads. Finally, the method is expected to have some appeal to practitioners as it is easy to use and understand.

A valid criticism of the method might be that distance scores never reach 1 (but a maximum m < 1) indicating total dissimilarity. Also, non-technical users, if they want to make use of the fuzzy predicate format, require to obtain some knowledge on the basics of fuzzy sets.

References

[Baldwin *et al,* 1995] J.F. Baldwin, T.P. Martin, B.W. Pilsworth, *Fril—Fuzzy and Evidential Reasoning in Artificial Intelligence,* Research Studies Pr., 1995.

[Buckles *et al,* 1982] B.P. Buckles, F.E. Petry, "A Fuzzy Model for Relational Databases", in *Fuzzy Sets and Systems,7 :213-226,* 1982.

[Chen and Hwang, 1992] S-J. Chen, C-L. Hwang, *Fuzzy Multiple Attribute Decision Making, Methods and Applications,* Springer Verlag, 1992.

[Dubitzky *et al,* 1996a] W. Dubitzky, J.G. Hughes, D.A. Bell, "Multiple Opinion Case-Based Reasoning and the Theory of Evidence", in *6th Int. Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems,* 447-452, 1996.

[Dubitzky *et al,* 1996b] W. Dubitzky, J.G. Hughes, D.A. Bell, "Case Memory and the Behaviouristic Model of Concepts", in *Advances in Case-Based Reasoning, 3rd European Workshop, EWCBR-96,* 120-134, 1996.

[Dutta *et al,* 1991] S. Dutta, P.P. Bonissone, "Integrating Case Based and Rule Based Reasoning: the Possibilistic Connection", in *Uncertainty in Artificial Intelligence 6,* 281-290, N.-Holland, 1991.

[George *et al,* 1993] R. George, B.P. Buckles, F.E. Petry, "Modelling Class Hierarchies in the Fuzzy Object-Oriented Data Model", in *Fuzzy Sets and Systems,* 60(3):259-272, 1993.

[Klir *et al,* 1988] G.J. Klir, T.A. Folger, *Fuzzy Sets, Uncertainty and Information,* Prentice Hall, Englewood Cliffs, NJ, 1988

[Petry, 1996] F.E. Petry, *Fuzzy Databases: Principles and Applications,* Kluver Academic Pub., MA, 1996.

[Tirri *et al,* 1996] H. Tirri, P. Kontkanen, P. Myllymaki, "A Bayesian Framework for Case-Based Reasoning", in *Advances in Case-Based Reasoning, 3rd European Workshop, EWCBR-96,* 113-127, 1996.

[Zadeh, 1973] L.A. Zadeh, "Outline of a New Approach to the Analysis of Complex Systems and Decision Processes", in *IEEE Transactions on Systems, Man, and Cybernetics,* 3(1), 28-44, 1973.