Evaluation in the Possibilistic Framework for Object Matching

Antoon Bronselaer, Axel Hallez, Guy De Tré
Department of Telecommunications and Information Processing
Sint-Pietersnieuwstraat 41, 9000 Ghent, Belgium
antoon.bronselaer@ugent.be, axel.hallez@ugent.be, guy.detre@ugent.be

Abstract

In current research, a possibilistic, hierarchical approach for identification of co-referent objects (also called object matching) has been proposed as a generalisation of the record linkage problem. This approach offers a natural view to the matching of co-referent objects and uses logical possibilistic operators to calculate a final result. The development of evaluation operators, to deliver basic possibilistic statements is, up till now, still in it's infancy. This paper introduces a formal basic definition of such operators and presents data type specific evaluation operators for sets.

Keywords: Possibilistic truth values, object co-reference, evaluators

1 Introduction

Detection of duplicate objects, for example database records or XML-documents, has been the source of many research in the past decades. In applications of data(base) merging, it is of vital importance to avoid duplicate storage and inconsistencies, which both imply inefficiency. In recent work, a hierarchical framework for object matching was introduced [6], which is a generic framework in the sense that the domain in which results are expressed, is left unspecified. In [2], a similar idea is introduced by defining *hierarchical fuzzy sets*, which are fuzzy sets defined over a subset of the elements of a finite hierarchy partially ordered by the "kind of" relation. The framework in [6] exploits the existence of an implicit entity structure, even when this structure is not reflected in the object storage model (e.g.: employee records are stored flat, but one can identify 'name' as a subobject, which in turn is constructed from a first and last name). Based on this framework, a possibilistic model for object matching can be constructed, which is further elaborated in [1] where the focus is mainly on aggregation and preference modeling. Such a model requires possibilistic evaluation operators for low-level comparison, i.e. attribute comparison. The development of such operators within the mentioned framework is still not well developed. Therefore, this work offers a prototype for such attribute evaluation operators in a possibilistic setting and illustrates it for the comparison of sets. The paper is structured as follows. In Section 2, some basic concepts are introduced. Section 3 defines a general possibilistic evaluation operator and summarizes interesting properties, which leads to an evaluation operator for sets in Section 4. Finally, the main contributions of this work are summarized in Section 5.

2 Preliminaries

A possibilistic truth value (PTV) is a possibility distribution (i.e. a fuzzy set) defined over the set of boolean values $I = \{T, F\}$, where T represents true and F represents false [9, 12]. They are used to express the uncer-

L. Magdalena, M. Ojeda-Aciego, J.L. Verdegay (eds): Proceedings of IPMU'08, pp. 1701–1708
Torremolinos (Málaga), June 22–27, 2008

tainty about the boolean value of a proposition. Let P denote the set of all propositions, then each $p \in P$ can be associated with a PTV $\tilde{p}$:

$$\tilde{p} = \{(T, \mu_{\tilde{p}}(T)), (F, \mu_{\tilde{p}}(F))\}$$

where $\mu_{\tilde{p}}(T)$ represents the possibility that p is true and $\mu_{\tilde{p}}(F)$ represents the possibility that p is false. The set of all PTVs is denoted by $\tilde{\wp}(I)$. In this work it is assumed that $\max(\mu_{\tilde{p}}(T), \mu_{\tilde{p}}(F)) = 1$, which supports the assumption that the universe $\{T, F\}$ is large enough to express the truth value of any proposition. It is possible to define generalizations $\tilde{R}$ of order relations R as follows:

$$\tilde{p}_1 \; \tilde{R} \; \tilde{p}_2 \Leftrightarrow$$

$$\begin{cases} \mu_{\tilde{p}_2}(F) \; R \; \mu_{\tilde{p}_1}(F), & \mu_{\tilde{p}_1}(T) = \mu_{\tilde{p}_2}(T) = 1 \\ \mu_{\tilde{p}_1}(T) \; R \; \mu_{\tilde{p}_2}(T), & \text{else} \end{cases}$$

3 Possibilistic evaluation

As mentioned before, this paper contributes to a possibilistic approach for object matching. The term 'object' should be interpreted quite general in the sense that an object is an arbitrarily complex description of some entity. It is assumed here that such descriptions reflect the natural structure of entities in a hierarchical way, more specific by using a *tree structure*. The scope of this paper lies on objects that share such a predefined structure. However, even when the structure is predefined, an entity can still be described in different ways. Descriptions that refer to the same entity are called *co-referent*. The goal of object matching is to detect co-referent objects. A possibilistic solution for this problem is inferred as follows. Given two objects we have the following affirmative proposition p:

$$p = \text{“}o_1 \text{ and } o_2 \text{ are co-referent”}$$

which can be evaluated to a boolean value. Now as non-equal objects can be co-referent, there is an implicit uncertainty about the boolean value of p, which can be modeled by a possibilistic truth value (Section 2). Hence, the problem of object matching is to provide the membership grades of the PTV associated with proposition p. Such a calculation is obtained by comparing the sub-objects defined in the object structure shared by both objects. The most basic sub-objects are called *attributes* and comparing the values of n attributes results in n propositions p_i:

$$p_i = \text{“the } i^{th} \text{ attribute has co-referent values”}$$

These attributes are sometimes assumed to be atomic, but in this work this assumption is omitted for the sake of generality. For example, when storing data on employees, the set of languages spoken by an employee is a non-atomic attribute. The operators that formulate possibilistic statements about such propositions are called *possibilistic comparative evaluators* or *evaluators* for short. The statements provided by evaluators are combined by using logical *aggregation operators* for PTVs, exploiting the implicit structure of objects. The above introduction about the

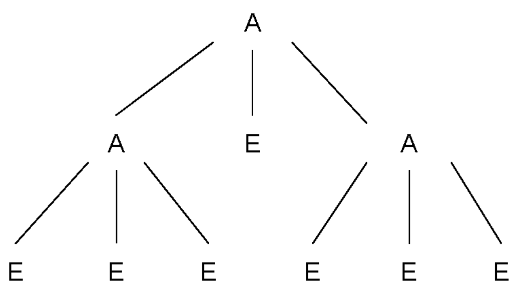

Figure 1: Possibilistic comparison scheme

hierarchical possibilistic framework for object matching is summerized in Figure 1, showing an example of a hierarchical possibilistic comparison scheme, where A is an aggregation operator for PTVs and E is an evaluation operator. Little work has been done on the development of evaluation operators that directly estimate the possibilities of the boolean value of p_i. For that purpose, we first define a generic form of such an operator.

Definition 1 (Evaluator)**.** Assume a universe U. For each couple of values $(u, u') \in U^2$ we can state an affirmative proposition p:

$$p = \text{“u and u' are co-referent”}$$

The uncertainty about the boolean value of p is given by a possibilistic evaluator for U, formally defined as:

$$E_U : U^2 \rightarrow \tilde{\wp}(I) : (u, u') \mapsto E_U(u, u')$$

with

$$E_U(u, u') = \{(T, \mu_{\tilde{p}}(T)), (F, \mu_{\tilde{p}}(F))\}$$

and

$$E_U(u, u') = E_U(u', u)$$

It was chosen to limit the number of axioms in Definition 1 to keep the class of evaluators as general as possible. However, symmetry is axiomatically required because uncertainty concerning equality of entities should always be symmetrical from a semantical point of view. Some examples of evaluators are:

$$E_{\mathcal{L}}(\{Dutch\}, \{French, Danish\}) = \{(F, 1)\}$$

$$E_{\mathcal{S}}(\text{“}J\ Lennon'', \text{“}Lennon\ John'') = \{(T, 1)\}$$

$$E_{I\!R}(37.5, 37.9) = \{(T, 1)(F, 0.2)\}$$

where $\mathcal{L}$ represents the set of all possible languages combinations, $\mathcal{S}$ the set of strings and $I\!R$ the set of real numbers. Looking at the examples, it becomes clear that, depending on the specific application, it is interesting to specify additional properties for evaluators. If one uses the equality relation on U, the following semantical constraint must be satisfied:

$$\forall (u, u') \in U^2 : u = u' \Rightarrow E_U(u, u') = \{(T, 1)\}$$

This condition, called *reflexivity*, states that when two values from U are equal, they are co-referent. A more strict constraint would be to assume that as soon as two values are not equal, it is to some extent possible that they are not co-referent:

$$\forall (u, u') \in U^2 : u = u' \Leftrightarrow E_U(u, u') = \{(T, 1)\}$$

This latter property is called *strong reflexivity*. A clear distinction is made between both properties, which both have their use in practical applications. In the scope of this paper, strong reflexivity is preferred because non-equality always introduces some uncertainty. Reflexivity describes a semantical connection between equality of values and their co-reference. However, the connection between difference of values and their co-reference is left unspecified up till now. One could argue that, in a strict sense, as soon as two values are different under the equality relation, the possibility that they are not co-referent should be 1 and the uncertainty should be reflected in the extent to which it is possible that the values are co-referent. Following such a reasoning leads to a very conservative approach. However, possibilities should not be interpreted absolute, but referential, which is why the following approach is more suitable. Initially we do not know anything about the co-reference of two values. This complete uncertainty is modeled by $\tilde{p} = \{(T, 1), (F, 1)\}$. Next, evidence is searched about the co-reference and if it is more possible that the values are co-referent than the opposite case, $\mu_{\tilde{p}}(F)$ should be smaller than 1. So the framework of PTVs for object matching states that when two values are not equal under some predefined equality relation, it still might be *more possible* that the values are co-referent. In the following, it is assumed that the evaluators are strong reflexive.

In some cases, it is possible to identify relations between propositions concerning co-reference. For instance, assume we have a proposition stating the co-reference of a and b, say $p_{a,b}$ and a proposition stating the same about b and c, say $p_{b,c}$. The uncertainty about the boolean truth values of these propositions is given by the PTVs $\tilde{p}_{a,b}$ and $\tilde{p}_{b,c}$. An interesting problem is what we know about $p_{a,c}$, the proposition stating that a and c are co-referent. In literature concerning similarity metrics, these relations are often referred to as *transitivity*, which is an implicit property of a similarity metric. In the possibilistic model for co-reference, such a direct transitivity is not present. Nevertheless, for a given evaluator, it might be possible and useful to derive a *conditional* possibility distribution over the domain $I = \{T, F\}$, say $\tilde{p}_{a,c}|\tilde{p}_{a,b}, \tilde{p}_{b,c}$ repre-

Table 1: Examples of conditional PTVs

$\tilde{p}_{a,b}$	$\tilde{p}_{b,c}$	$\tilde{p}_{a,c} \vert \tilde{p}_{a,b}, \tilde{p}_{b,c}$
{(T,1)}	{(T,1)}	{(T,1)}
{(T,1)}	{(F,1)}	{(F,1)}
{(F,1)}	{(T,1)}	{(F,1)}
{((F,1)}	{(F,1)}	{(T,1),(F,1)}
{(T,1)}	{(T,1),(F,1)}	{(T,1),(F,1)}
{(F,1)}	{(T,1),(F,1)}	{(T,1),(F,1)}
{(T,1)}	{(T,1),(F,0.1)}	{(T,1),(F,0.1)}
{(T,1),(F,0.3)}	{(T,1),(F,0.1)}	{(T,1),(F,0.3)}
{(T,0.5),(F,1)}	{(T,1),(F,0.1)}	{(T,0.5),(F,1)}
{(T,0.5),(F,1)}	{(T,0.3),(F,1)}	{(T,1),(F,1)}
{(T,1),(F,0.9)}	{(T,1)}	{(T,1),(F,0.9)}
{(T,1),(F,1)}	{(T,1),(F,1)}	{(T,1),(F,1)}

senting the uncertainty about the boolean value of $p_{a,c}$, given $\tilde{p}_{a,b}$ and $\tilde{p}_{b,c}$. Let us begin by describing the relations that we have. If we have a (strong) indication that a and b are co-referent and (strong) indication that b and c are co-referent, then both indications combined are an indication for the co-reference of a and c. Now, an indication that a and b (resp. b and c) are co-referent combined with an indication that b and c (resp. a and b) are *not* co-referent, yields an indication that a and c are not co-referent. Finally, an indication that a and b are not co-referent combined with an indication that b and c are not co-referent, tells us nothing about the co-reference of a and c. As an indicative measure we consider necessity, which reflects *certainty* rather than possibility and is derived as follows:

$$Nec_p(T) = 1 - \mu_{\tilde{p}}(F)$$
$$Nec_p(F) = 1 - \mu_{\tilde{p}}(T)$$

Based on these transformations and the following notations of conditional necessity:

$$\mathcal{N}_{p_{a,c}}(T) = \left(Nec_{p_{a,c}}(T) | \tilde{p}_{a,b}, \tilde{p}_{b,c}\right)$$
$$\mathcal{N}_{p_{a,c}}(F) = \left(Nec_{p_{a,c}}(F) | \tilde{p}_{a,b}, \tilde{p}_{b,c}\right)$$

the above descriptions of the (un)certainty relations between propositions are formalised:

$$\mathcal{N}_{p_{a,c}}(T) \geq \left(Nec_{p_{a,b}}(T) \wedge Nec_{p_{b,c}}(T)\right)$$
$$\mathcal{N}_{p_{a,c}}(F) \geq \left(Nec_{p_{a,b}}(T) \wedge Nec_{p_{b,c}}(F)\right) \vee \left(Nec_{p_{a,b}}(F) \wedge Nec_{p_{b,c}}(T)\right)$$

where the conjunction operator $\wedge$ is min and the disjunction operator $\vee$ is max. By adding the normalisation condition of necessities $\mathcal{N}_{p_{a,c}}(T) \cdot \mathcal{N}_{p_{a,c}}(F) = 0$, we can determine the conditional necessities and hence the conditional possibilities by using the inverse transformations. Table 3 contains some examples of derived conditional possibilistic truth values. The examples show how uncertainty about the basic propositions is contained in the conditional distribution. When there are indications that both basic properties are false, the conditional distribution will reflect complete uncertainty, just as is required. The conditional PTV provides an upper bound for the uncertainty about proposition $p_{a,c}$, meaning that if additional information about this proposition becomes available, the resulting uncertainty must be smaller than or equal to the conditional uncertainty we had before the addition of information. However, adding new information that results in strictly less uncertainty than the conditional uncertainty, but implies an indication toward a different truth value signifies a contradiction. The occurrence of such contradictions depends on both the evaluator and the specific problem domain. For example, when the evaluator is not strong reflexive and we compare sets, such contradictions can occur. An evaluator for which the inferred conditional possibility distribution is never in contradiction with additional information delivered by the evaluator, is called a *consistent* evaluator. In other words, using a consistent evaluator E implies that $\tilde{p}_{a,c}|\tilde{p}_{a,b}, \tilde{p}_{b,c}$ and $E(a, c)$ both agree on the truth value that is most possible. The determination of a conditional possibility distribution has important practical applications, for example with a highly complex evaluator used in an object matching environment, it can save computational resources. As a final remark to conclude this section, note that both estimators of possibility are completely independent of any notion of similarity. Possibilities can be linked to similarities in the sense that when values are very similar, the possibility that they are co-referent should be high. Nevertheless one should always be careful not to confuse possibilities with degrees of (dis)similarity. In what follows, Definition 1 will be specified for the set data type, thereby using some principles introduced by Dubois and Prade on set comparison [5].

4 Set evaluation

As mentioned in Section 3 attributes can be non-atomic and sets are a first example to illustrate the relevance of this assumption. The set datatype is used on attribute level in different situations, for example to denote a collection of items or to model the seperate words in a sentence such as done in several string comparison systems [8]. It is explicitly stated that when set comparison is discussed, we do not refer to a set of constraints (which is actually the topic of object matching), but rather some separate collection of items for which the co-referential uncertainty should be estimated. In the following, two approaches for set evaluation, respectively *hard set evaluation* and *soft set evaluation*, are discussed.

4.1 Hard set evaluation

The first approach for set evaluation extends some regular comparison techniques for sets to $\tilde{\wp}(I)$. These comparison strategies calculate a result based on (well known) set functions [3]. After deriving sets by using such functions, an important step is the mapping of the derived sets to the unit interval. Dubois and Prade use *fuzzy measures* in this step of the comparison ([5]). Such measures are defined as follows. Assume a universe U and two subsets of U, A and B. A *fuzzy measure* [10] is a mapping from $\wp(U)$ to $[0, 1]$ satisfying:

$$\gamma(\emptyset) = 0$$
$$\gamma(U) = 1$$
$$A \subset B \Rightarrow \gamma(A) \leq \gamma(B)$$

Our approach requires an estimation of the possibilities that two given sets are (not) co-referent. To do so, a couple of bipolar fuzzy measures is used to make a distinction between positive and negative information delivered by the results of set functions. In the context of sets, the positive information is contained in the elements shared by the sets and the negative information is contained by the elements that do not occur in both sets. Hence, a formal way of defining a set evaluator is:

Definition 2 (Hard Set Evaluator)**.** Assume a universe U. A hard set evaluator $E^h_{\wp(U)}$ is an evaluator as defined in Definition 1 where:

$$\mu_{\tilde{p}}(T) = s\frac{\gamma^T(A \cap B)}{\gamma^T(A \cup B)}$$
$$\mu_{\tilde{p}}(F) = s\frac{\gamma^F(A\Delta B)}{\gamma^F(A \cup B)}$$

where $A\Delta B = (A\cup B)\cap(\overline{A}\cup\overline{B})$ is the symmetrical difference of two sets and s is a scaling factor to ensure normality.

The fuzzy measures evaluate the relevance of the elements in a set. The estimation for T expresses the ratio of relevance of elements in the intersection and in the union. Similarly, the estimation for F expresses the ratio of relevance of elements in the symmetrical difference and the union. The relevance of an element being in the intersection might differ strongly from the relevance of that element being in the symmetric difference. More specific, $x \in (A \cap B)$ might have low relevance for the possibility that A and B are co-referent, while $x \in (A\Delta B)$ might be very relevant for the possibility that A and B are not co-referent. For both γ's, a simple example is:

$$\gamma(A) = \frac{|A|}{|U|}$$

4.2 Soft set evaluation

The approach presented above is based on set functions that use an equality relation '=' on the universe of discourse. A more flexible approach avoids the strict equality of elements. Instead, it states that elements themselves can be co-referent without being equal. More specific, the evaluation of sets relies on lower level evaluations of it's elements. Using such a low-level evaluator produces a sequence of PTVs representing uncertainty about boolean values of propositions concerning the co-reference of elements from U. Aggregation of these PTVs results in a single PTV representing the uncertainty about the co-reference of the sets. In what follows, it is assumed that an evaluator is present to estimate the uncertainty about propositions like "a and b are co-referent" with $a, b \in U$. This evaluator is referred to as E_U and must satisfy Definition 1 and the assumptions made in Section 3, more specific the assumption of strong reflexivity. Assume next two subsets of U, say A and B with $|A| \leq |B|$. The key idea is to create an injective mapping ι from element of A to elements of B. During the construction of ι, element couples that are more possible to be co-referent are given higher preference to be element of ι. Elements that are mapped to each other are used to create affirmative propositions about co-reference on element level. The procedure can be split up into some basic steps. First, Algorithm 1 creates a matrix of PTVs expressing uncertainty about the co-reference of elements from A and B that are not equal. The intersection $A \cap B$ is treated separately because, due to the strong reflexivity, equal elements are certainly co-referent, which is why Algorithm 1 maps the elements of $A \cap B$ to themselves. The functions $r(.)$ and $c(.)$ provide

Algorithm 1 Matrix generation

Require: $A, B \subset U \wedge |A| \leq |B|$
Ensure: A matrix M of PTVs
$C \leftarrow A \cap B$
$\forall x \in C : \iota(x) = x$
$A \leftarrow A \backslash C$
$B \leftarrow B \backslash C$
$\forall a \in A, b \in B : M[r(a), c(b)] = E_U(a, b)$

one-to-one mappings of elements from A and B to row and column indexes which are natural numbers. Note that in Algorithm 1 the variables A and B are overwritten. Hence, in what follows it is assumed that, after execution of Algorithm 1, $A \cap B = \emptyset$, which simplifies our notations. Having the matrix M, we want to iteratively find the largest PTVs, add it's location to the mapping and then remove the row and column of that location. This process is equivalent to Algorithm 2 which

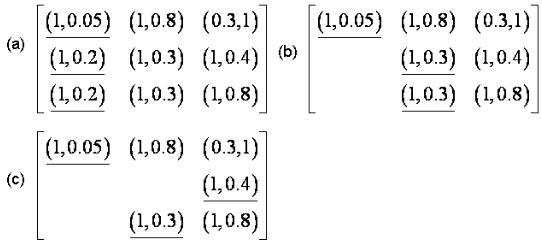

Figure 2: Example of element mapping

is explained as follows. For each row in M, the largest PTV is located with the understanding that comparison of PTVs is based on generalized order relations as explained in Section 2. If two rows, say r_1 and r_2 exist with the same location of the largest PTV, a *conflict* is present. These conflicts are resolved one at a time as follows. If the PTVs are different, a subprocedure called **search** disables the position of the current maximum on the row with the smallest current maximum and searches a new maximum for that row. In doing so, disabled maxima are not taken into account. If the PTVs are equal on both rows, a subprocedure called **choose** will identify which row should be passed to procedure **search** for relocation of it's maximum. The procedure selects the remaining PTVs from each row (i.e. enabled positions on the row) which results in two multisets of PTVs, say M_1 and M_2. Now, if $M_1 = M_2$, both rows contain the same PTVs on enabled positions. By convention, in this case we choose r_1. If $M_1 \subset M_2$ or $M_2 \subset M_1$, obviously the row corresponding with largest multiset is chosen because it contains all information captured by the smaller multiset. If neither of these cases yield, we subtract $M_1 \cap M_2$ from both multisets and the multiset containing the largest PTV after subtraction is chosen. This way the largest possible PTVs are left for future maximum relocation. An example of the element mapping described by Algorithm 2 is shown in Figure 2. For simplification, the PTVs are shown in a shorter notation with first the possibility of T and second the possibility of F. Hence, $\tilde{p} = \{(T, \mu_{\tilde{p}}(T)), (F, \mu_{\tilde{p}}(F))\}$ is represented as $(\mu_{\tilde{p}}(T), \mu_{\tilde{p}}(F))$. Further on, the current maxima are marked as the location where the PTV is underlined and when a position is disabled, the corresponding PTV is deleted. From step (a) to (b) the conflicts between the three rows are resolved. Because row 1 contains the largest PTV, the maxima on row 2 and 3 are relocated. From step (b) to (c) the conflict between row 2 and 3 is resolved on column 2. Both PTVs are equal so **choose** will select row 2 for maximum relocation because $(1, 0.4)\tilde{>}(1, 0.8)$. In step (c) no conflicts occur and the algorithm stops. Algorithm 2 produces the final injective mapping ι from elements of the smallest set (A) to the largest set (B). Each couple of elements from the final mapping is linked with one PTV expressing uncertainty about their co-reference, which induces a sequence *seq* of PTVs. In addition, $|B| - |A|$ elements from B will not occur in the mapping, so $\{(F, 1)\}$ should be added $|B| - |A|$ times to *seq*. The generated sequence *seq* represents the uncertainty about the co-reference of elements form B with elements from A. The final step requires an aggregation operator, such as (weighted) conjunctive and disjunctive operators for PTVs [4, 7], to infer one PTV expressing uncertainty about the co-reference of the sets. When using such operators it should be emphasized that the algorithm provided ensures only uniqueness of *seq*, not the uniqueness of the mapping of elements implying this sequence. Hence, if there is a difference in preference amongst elements, the element mapping itself becomes important and can be chosen to optimize a predefined criterion. However, such a mapping optimisation is outside the scope of this paper. The foregoing discussions lead to the following definition of soft set evaluation.

Algorithm 2 Element mapping

Require: $(|A| \times |B|)$-matrix M of PTVs
Ensure: Injective mapping ι
  for all $a \in A$ **do**
    $m[r(a)] \leftarrow \arg\max_{b \in B} M[r(a), c(b)]$
  end for
  while $\exists x \neq y \wedge m[r(x)] = m[r(y)]$ **do**
    $\tilde{p}_1 \leftarrow M[r(x)][c(m[r(x)])]$
    $\tilde{p}_2 \leftarrow M[r(y)][c(m[r(y)])]$
    if $\tilde{p}_1 = \tilde{p}_2$ **then**
      $d \leftarrow$**choose**$(M[r(x)], M[r(y)])$
    end if
    if $\tilde{p}_1 \tilde{<} \tilde{p}_2 \vee d = r(x)$ **then**
      $m[r(x)] \leftarrow$**search**$(M[r(x)])$
    else
      $m[r(y)] \leftarrow$**search**$(M[r(y)])$
    end if
  end while
  $\forall a \in A : \iota(a) = c^{-1}(m[r(a)])$

Definition 3 (Soft Set Evaluator)**.** Assume a universe U and two subsets A and B with $|A| \leq |B|$. The soft set evaluator is an evaluator satisfying:

$$E^s_{\wp(U)}(A, B) = F(\tilde{p}_1, ..., \tilde{p}_{|B|})$$

where the first $|A|$ PTVs are produced under ι and E_U and the last $|B| - |A|$ PTVs are $\{F, 1\}$. F is an arbitrary aggregation function for PTVs.

The relevance of investigating set evaluation is emphasized due to it's important applications in for example multiset evaluation and string evaluation. Both topics will be investigated in future research.

5 Conclusion

A generalization of the record linkage problem know as the object matching problem, has been tackled from a possibilistic point of view in current research. In order to further elaborate this model, we have introduced a formal definition of so called evaluation operators in the domain of possibilistic truth values, which estimate possibilities concerning co-reference on a low level. As an application, we have presented evaluation operators for sets. A first approach is an extension of past research on set comparison, while the second approach benefits from the implicit non-atomicity of sets.

References

[1] A. Bronselaer, G. De Tré, A. Hallez (2007). Dynamic preference modeling in flexible object matching. In *Proceedings of the Eurofuse Workshop on New Trends in Preference Modeling*, p. 191-195, Jaén, Spain, 2007.

[2] R. Thomopoulos, P. Buche, O. Haemmerlé (2006). Fuzzy sets defined on a hierarchical domain. In *IEEE Transactions on Knowledge and Data Engineering*, p. 1397-1410, 18(10), .

[3] V. Cross, T. Sudkamp (2002). Similarity and Compatibility in Fuzzy Set Theory: assessment and applications. In *Physica-Verlag*, 2002

[4] G. De Cooman (1995). Towards a possibilistic logic. In *Fuzzy Set Theory and Advanced Mathematical Applications*, p. 89–133, 1995.

[5] D. Dubois, H. Prade (1982). A unifying view of comparison indices in a fuzzy set-theoretic framework. In *Recent Developments in Fuzzy Set and Possibility Theory*, p. 3-13, 1982.

[6] A. Hallez, G. De Tré (2007). A Hierarchical approach to object comparison. In *Proceedings of IFSA World Congres on Foundations of fuzzy logic and soft computing*, p. 191-198, Cancun, Mexico, 2007.

[7] T. Matthé, G. De Tré, A. Hallez (2007). Weighted Conjunctive and Disjunctive Aggregation of Possibilistic Truth Values. In *Proceedings of IFSA World Congres on Foundations of fuzzy logic and soft computing*, p. 171-180, Cancun, Mexico, 2007.

[8] A. Monge, C. Elkan (1996). The Field Matching Problem: Algorithms and Applications. In *Proceedings of the International Conference on Knowledge Discovery and Data Mining*, p. 267-270.

[9] H. Prade (1982). Possibility sets, fuzzy sets and their relation to Lukasiewicz logic. In *Proc 12th Int Symp on Multiple-Valued Logic*, p. 223-227, 1982.

[10] M. Sugeno (1974). Theory of fuzzy integrals and it's applications. In *PhD thesis*, Tokyo Institute of Technology, 1974.

[11] L. Zadeh (1965). Fuzzy sets. In *Information and Control*, vol. 8, no. 3, p. 338-353, 1965.

[12] L. Zadeh, "Fuzzy sets as a basis for a theory of possibility", *Fuzzy Sets and Systems*, p. 3-28, 1978.