Using Generalized Combination Measure from Dombi and Yager type of $T$-norms and $T$-conorms in Classification

Kalle Saastamoinen and Jaakko Ketola
Department of Information Technology
Lappeenranta University of Technology
Lappeenranta, Finland
Email: kalle.saastamoinen@lut.fi

Abstract—In this paper we use a comparison measure which can be adapted to various different applications and has many good properties for further analysis. This comparison measure is based on the adaptive aggregation of the $t$-norms and $t$-conorms. We test this measure in classification tasks, with Dombi and Yager type of $t$-norm and $t$-conorm operators. The classification results are compared to plain $t$-norms, $t$-conorms, and Schweizer-Sklar-Lukasiewicz-based equivalence. It is shown that the presented measure is able to give highly competitive results in classification.

Index Terms—Fuzzy and neural systems, Compensation, Classification, Differential evolution, Many-valued connectives, Aggregation

I. INTRODUCTION

In all areas where human beings are involved, some kind of measures for comparison are needed. In soft computing these areas are for example classification, pattern recognition, clustering, expert systems, medical diagnosis systems, decision support systems, fuzzy control etc. Recent areas are for example web-search engines, where information retrieval is of high importance. Many times the measures are intuitively taken, which can be a crucial mistake and can lead to so called black-box systems which can work, but no one is able to say how and why they work. This is the case for example when we use neural networks [1]. In fact the first thing to be considered in systems should be used measures.

In comparison tasks, such as classification, it is often very important to be able to model linguistic data for computers. We can see that words like and and or act in significant roles in this transformation from a comparison model to a computer. Unions and intersections can be seen as set theoretical correspondences of these words. However, if we use traditional unions and intersections we model only the situation whether or not a given object is a member of a class. In most comparison tasks the question is not this simple. In reality, the important thing is the degree to which the objects belong to the classes [2]. This degree can be achieved by the use of binary relations in $[0,1]$ called $t$-norms and $t$-conorms. These are used for example in the fields of statistical metric spaces as a tool for generalizing the classical triangular inequality [3], [4], which has been successfully used in expert knowledge systems since 1976 [5] and in many other fields.

$T$-norms are widely accepted as equivalent to many-valued intersections, and correspondingly, $t$-conorms as equivalent to many-valued unions. Semantically, these two sets of norms can be understood as generalizations of two-valued logical operators AND and OR or conjunction and disjunction connectives, respectively. In practice it is known that different choices of these operators lead to radically different results [6]. The $t$-norm gives minimum type of compensation, while the $t$-conorm gives maximum type of compensation. This means that $t$-norms tend to give more value for the small values, while $t$-conorms give more value for the big values in the intervals they are used. In practice neither of these connectives fit the collected data appropriately. There is still a lot of information that is left between them. An important issue when dealing with $t$-norms and $t$-conorms is the question of how to combine them in a meaningful way, since neither of these connectives alone gives general compensation for the values they are adapted to. For this reason we should use a measure that compensates for this gap of values between them. We have used the generalized mean in this paper as our compensation operator. The way the generalized mean works for compensation can be found in article [7]. The full scope of some well-known aggregation operators are demonstrated in figure 1.

Fig. 1. Compensation of $t$-norms and $t$-conorms

In paper [8] we have defined measures based on the use of the generalized mean, weights, $t$-norms and $t$-conorms. Here we will show a measure which is a combination of the $t$-norm and $t$-conorm. This measure uses adaptive aggregation in the same way as presented in [9] and [10].

The motivation for this article is to show that the results achieved by the use of $t$-norms and $t$-conorms can get significantly better if we combine these two norms and use the
resulting compensatory measure instead. We will also show that the results are very good even when they are compared, versus the results that can be achieved using equivalence relations. We have chosen to use classification as our test bench for these measures, since due the close connection between real world classification problems and many-valued logic, it is clear that the study of connectives used for comparison may shed some light on the general problems of decision-making [12].

II. COMBINATION OF MANY-VALUED INTERSECTIONS AND UNIONS

We have defined in papers [9] and [10] an operator named Generalized Weighted Norm Operator (GWNO), which is used to compensate many-valued intersections and unions. The idea is to compensate values that unions and intersections alone are capable to give, as shown in figure 1. The measure itself can be defined as shown in equation 1. The relations in this paper are of the form \( R : ([0, 1]^n)^2 \to [0, 1] \). The tested vectors \( x(i), y(i) \in [0, 1], \forall i \in \mathbb{N} \), mean values \( m \in \mathbb{R} \) and \( m \neq 0 \) and all the weights \( w_i, \omega_c, \) and \( \omega_d \) hold that \( w_i \geq 0 \) with \( \sum_{i=1}^{n} w_i = 1 \), and \( p \) is a parameter value, which corresponds to the comparison measure class.

Definition 1: Compensative measure based on the t-norm and t-conorm with generalized mean and weights:

\[
F_p(x(i), y(i)) = \left( \sum_{i=1}^{n} \left( w_i c_i^p (x(i) + y(i)) + (1 - w_i) (d_i^p (x(i) + y(i))) \right)^m \right)^{\frac{1}{m}},
\]

where \( i = 1, \ldots, n \), \( p \) is a parameter combined to the comparison measure class.

Definition 2: Measure based on Dombi [13] class of t-norm with generalized mean and weights:

\[
C_D(x(i), y(i)) = \left( \sum_{i=1}^{n} C_{\omega_c} \left( \left( \frac{1}{x(i)} - 1 \right)^p + \left( \frac{1}{y(i)} - 1 \right)^p \right)^{-m} \right)^{\frac{1}{m}},
\]

where \( p \geq 0 \) and \( i = 1, \ldots, n \).

Definition 3: Measure based on Yager [14] class of t-norm with generalized mean and weights:

\[
C_Y(x(i), y(i)) = \left( \sum_{i=1}^{n} C_{\omega_c} \left( 1 - \min \left( (x(i))^p + (y(i))^p \right) \right)^{m} \right)^{\frac{1}{m}},
\]

where \( p \geq 0 \) and \( i = 1, \ldots, n \).

Definition 4: Measure based on Dombi [13] class of many valued union with generalized mean and weights:

\[
D_D(x(i), y(i)) = \left( \sum_{i=1}^{n} \omega_d \left( 1 + \left( \frac{1}{x(i)} - 1 \right)^-p + \left( \frac{1}{y(i)} - 1 \right)^-p \right)^{-m} \right)^{\frac{1}{m}},
\]

where \( p > 0 \) and \( i = 1, \ldots, n \).

Definition 5: Measure based on Yager [14] class of t-conorm with generalized mean and weights:

\[
D_Y(x(i), y(i)) = \left( \sum_{i=1}^{n} \omega_d \left( \min \left\{ 1, (x(i))^p + (y(i))^{p} \right\} \right)^{m} \right)^{\frac{1}{m}},
\]

where \( p > 0 \) and \( i = 1, \ldots, n \).

A. Schweizer-Sklar-Lukasiewicz-Based Equivalence Measure

Here we present a measure which rises from the functional definition for the implications given in [15] and from the use of an algebraic bounded product [11] as the many-valued conjunction for combining implications. We note that Lukasiewicz [16] and Schweizer and Sklar [17] types of implications form almost the same many-valued equivalence measures when these equivalences are formed by using the algebraic bounded product to combine corresponding implications. The only difference between these measures is that while Lukasiewicz is defined for positive parameter values of \( p \), Schweizer & Sklar is defined also for negative parameter values of \( p \). Therefore, these two equivalences are easily combined by taking the parameter values which go from negative to positive, so \( p \in [-\infty, \infty] \). When we combine the equivalence weights \( w_i \in [0, 1], \forall i \in \mathbb{N} \) and the generalized mean we reach

Definition 6: Generalized logical-functional weighted equivalence based on Schweizer-Sklar-Lukasiewicz:

\[
E_{SSL}(x(i), y(i)) = \left( \sum_{i=1}^{n} w_i \left( 1 - |x(i) - y(i)|^p \right) \right)^{\frac{1}{p}},
\]

where \( p > 0 \) and \( i = 1, \ldots, n \).

III. CLASSIFICATION

We have used classification as the test bench for our comparison measures in this paper.

A. Datasets

For our classification test, we selected two different data sets available in [18]. The data sets chosen for the test were Ionosphere and Postoperative. These sets differ greatly in the magnitude of instances and the number of predictive attribute values.

**Ionosphere:** This is radar data, where the targets are free electrons in the ionosphere. There are two classes: “Good” and “Bad”. “Good” radar returns are those showing evidence of some type of structure in the ionosphere. “Bad” returns are those that do not; their signals pass through the ionosphere. The number of instances is 351. The number of attributes is 34 plus the class attribute.

**Postoperative:** This is medical dataset with some logistic aspect. The classification task of this database is to determine where patients in a postoperative recovery area should be sent to next. There are three classes corresponding to the recovery area. The number of instances is 90. The number of attributes is 9, including the class attribute. Attribute 8 has 3 missing values, which we have replaced by the mean of the remaining values of attribute 8.
B. Classifier

Our goal is to classify objects, each characterized by one feature vector in \([0,1]^n\), into different classes. The assumption that the vectors belong to \([0,1]^n\) is not restrictive since the appropriate shift and normalization can be done for any space \([a,b]^m\). Equations 1, 2, 4, 3, 5 and 6 can be used to compare objects to classes. In the algorithm below we denote feature vectors by \(u_m, m = 1, \ldots, M\), i.e. we have \(M\) objects to be classified. We also denote the number of different classes by \(L\).

The general classification procedure has the following steps:
1) Choose the \(m\) and \(p\) used with the comparison measure.
2) Choose ideal vectors \(f_l\) that present the classes as well as possible. These ideal vectors can either be given by expert knowledge or calculated in some way from a training set. We have calculated one ideal vector \(f_l\) for each class \(l\) by using the generalized mean, i.e.

\[
f_l(i) = \left( \frac{1}{n_l} \sum_{j=1}^{n_l} (v_{j,l}(i))^m \right)^{\frac{1}{m}} \quad \forall i \in \{1, \ldots, n\}, \quad (7)
\]

where vectors \(v_{j,l}\) are known to belong to class \(l\) and \(n_l\) is the number of these vectors.
3) Choose values for weights \(w\). Here we can choose whether we will use random weights or differential evolution (DE) for finding the best possible weights.
4) Compare each feature vector \(u\) to each ideal vector \(f_l\), i.e. calculate the measure \(\text{measure}(u_m,f_l;m;p;w)\) for all \(m \in \{1, \ldots, M\}\) and \(l = 1, \ldots, L\). Here the word \(\text{measure}\) refers to any comparison measure we prefer to use.
5) Make the decision that the feature vector \(u_m\) belongs to that class \(k\) for which the \(\text{measure}(u_m,f_k;m;p;w) = \max\{\text{measure}(u_m,f_l;m;p;w) \mid l = 1, \ldots, L\}\).

In short, we select those members to belong to one class, which have maximal degree of being a member of that class. This is done after we have gone through all the classes. Therefore, this classification could be described as a maximal correspondence-based classification procedure.

IV. RESULTS

Here we will show the maximal classification results from the chosen datasets using the comparison measures defined in equations 1, 2, 4, 3, 5 and 6. We have used two separate strategies for finding the right weights. With randomized weights (RND) the weighting was done by randomly shooting weights 200 times into each tested \(m\) and \(p\) value. In our second strategy optimized weights were found by the use of differential evolution (DE).

Table I shows that in the classification of the Ionospheric data the compensated measure in equation 1 gives generally better results than the corresponding \(t\)-norms or \(t\)-conorms alone are able to give. The best result, 94.886% correct classification is reached by the use of a compensative measure in equation 1 from a Yager type intersection in equation 3 and union in equation 5 with differential evolution for weight optimization. The next best result, 90.909% correct classification is achieved by the use of a combination of Dombi measures in equation 2 and in equation 4 with differential evolutionary optimized weights. The same result, 90.909% is also achieved by the use of weighted equivalence based on Schweizer-Sklar-Lukasiewicz in equation 6 with randomly shot weights.

The best result, 82.222% correct classification for the Postoperative data can be reached by the use of the generalized Dombi type of \(t\)-conorm in equation 4 with DE for weight optimization or by the use of a compensated measure in equation 1 from Dombi or Yager type intersections and unions with randomly shot weights.

Figure 2 shows that all the measures tested with differential evolutionary optimized weights have huge areas of different \(p\)- and \(m\)-values where the best results achieved are quite stable. Especially huge stable areas give the Dombi \(t\)-norm in equation 2, compensative measure in equation 1 with Yager \(t\)-norm in equation 3 and \(t\)-conorm in equation 5, Schweizer-Sklar-Lukasiewicz in equation 6. We also see that when we combine Yager type intersection and union, the compensation effect of the comparison measure in equation 1 can be clearly seen from the results, and now actually the best result is achieved in the negative area of \(m\)-values. On the other hand, the results with the plain Yager \(t\)-norm in equation 3 or \(t\)-conorm in equation 5 are very bad in the negative area.

Figure 3 shows that all the measures tested with randomly shot weights have again huge areas of different \(p\)- and \(m\)-values where the best results achieved are very stable. The compensative effect of the comparison measure in equation 1 can clearly be seen in figure 3. The topologies of the results with compensated measures come now very close to the topology achieved by the Schweizer-Sklar-Lukasiewicz equivalence in equation 6.

V. CONCLUSIONS

In this article we have shown how compensation measure 1 can be used for classification. We showed that classification results normally get better if we use compensative measure 1 for \(t\)-norms and \(t\)-conorms, compared to the results that can be achieved by using these many-valued intersections and unions alone. We also showed that the results were better than the

### Table I

<table>
<thead>
<tr>
<th>Measure</th>
<th>Ionosphere</th>
<th>Postoperative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dombi (1_{DE})</td>
<td>0.89773</td>
<td>0.8</td>
</tr>
<tr>
<td>Dombi (tco_{DE})</td>
<td>0.88068</td>
<td>0.82222</td>
</tr>
<tr>
<td>Dombi (combo_{DE})</td>
<td>0.90909</td>
<td>0.8</td>
</tr>
<tr>
<td>Dombi (1_{RND})</td>
<td>0.85227</td>
<td>0.77778</td>
</tr>
<tr>
<td>Dombi (tco_{RND})</td>
<td>0.67614</td>
<td>0.82222</td>
</tr>
<tr>
<td>Dombi (combo_{RND})</td>
<td>0.88068</td>
<td>0.82222</td>
</tr>
<tr>
<td>Yager (IDE)</td>
<td>0.90341</td>
<td>0.8</td>
</tr>
<tr>
<td>Yager (tco_{IDE})</td>
<td>0.89773</td>
<td>0.8</td>
</tr>
<tr>
<td>Yager (combo_{IDE})</td>
<td>0.94886</td>
<td>0.8</td>
</tr>
<tr>
<td>Yager (1_{RND})</td>
<td>0.875</td>
<td>0.77778</td>
</tr>
<tr>
<td>Yager (tco_{RND})</td>
<td>0.86364</td>
<td>0.82222</td>
</tr>
<tr>
<td>Yager (combo_{RND})</td>
<td>0.86364</td>
<td>0.82222</td>
</tr>
<tr>
<td>SS (ekv_{DE})</td>
<td>0.89773</td>
<td>0.77778</td>
</tr>
<tr>
<td>SS (ekv_{RND})</td>
<td>0.90909</td>
<td>0.73333</td>
</tr>
</tbody>
</table>
results achieved by using many-valued equivalence relation 6, which is a notable result since many-valued equivalence relations are usually able to give very competitive results in classification tasks.

ACKNOWLEDGEMENTS

This work was partially supported by the East Finland Graduate School in Computer Science and Engineering (ECSE).

REFERENCES

[9] Saastamoinen K. and Ketola J., Generalized Weighted Norm Operator GWNO - Case Study: Classification, sent for reviews to the international journal.