Visualizing the computational intelligence field

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

Nees Jan van Eck, Ludo Waltman,
Jan van den Berg, and Uzay Kaymak
*Erasmus University Rotterdam,
THE NETHERLANDS*

Visualizing the Computational Intelligence Field

In this paper, we visualize the structure and the evolution of the computational intelligence (CI) field.[1] Based on our visualizations, we analyze the way in which the CI field is divided into several subfields. The visualizations provide insight into the characteristics of each subfield and into the relations between the subfields. By comparing two visualizations, one based on data from 2002 and one based on data from 2006, we examine how the CI field has evolved over the last years. A quantitative analysis of the data further identifies a number of emerging areas within the CI field.

The data that we use consist of the abstracts of the papers presented at the IEEE World Congress on Computational Intelligence (WCCI) in 2002 and 2006. Using a fully automatic procedure, so-called concept maps are constructed from the data. These maps visualize the associations between the main concepts in the CI field. Our analysis of the structure and the evolution of the CI field is largely based on the constructed concept maps.

Methodology

A concept map is a map in which concepts are located in such a way that the closeness of two concepts reflects their association as accurately as possible. We used the following procedure to construct concept maps of the CI field. First, for each of the abstracts of the papers presented at the WCCI 2002

and the WCCI 2006, the concepts occurring in the abstract were identified. This was done using a simple thesaurus of the CI field that we had constructed ourselves (for more details, see [1]). In the next step, co-occurrences of concepts were counted. Let there be n concepts, denoted by $1, \ldots, n$. The co-occurrence frequency c_{ij} of the concepts i and j is defined as the number of abstracts in which the concepts i and j both occur. Based on concepts' co-occurrence frequencies, the associations between pairs of concepts were calculated. The association between two concepts was quantified using a measure called association strength. The association strength a_{ij} of the concepts i and j is defined as

$$a_{ij} = \frac{m c_{ij}}{c_{ii} c_{jj}}, \quad \text{for } i \neq j \qquad (1)$$

where c_{ii} denotes the number of abstracts in which concept i occurs and m denotes the total number of abstracts. The association strength of two concepts can be interpreted as the estimated co-occurrence probability of the concepts normalized for the estimated co-occurrence probability obtained under the assumption that occurrences of the two concepts are statistically independent [1]. Concept maps of the CI field were constructed based on the association strengths of concepts. Association strengths rather than co-occurrence frequencies were used to construct concept maps in order to ensure that all concepts were treated in the same way regardless of the number of abstracts in which they occur. The con-

struction of concept maps was accomplished using a new method called VOS. This method will be introduced below. Separate concept maps were constructed for 2002 and 2006, and only two-dimensional maps were considered.

To display a concept map, we used a Java applet that we refer to as the concept map viewer. The concept map viewer indicates the location of a concept in a concept map by displaying a label at that location. This label shows a term that identifies the concept. The size of the label reflects the importance of the concept, which was measured by the number of abstracts in which the concept occurs. The concept map viewer has scroll, zoom, and search functionality to support a comprehensive examination of a concept map. In addition to the concept map viewer, we have also developed a computer program that displays so-called concept density maps. A concept density map is a concept map in which colors are used to indicate the density of concepts. Concept density maps are especially useful to get a quick overview of the various clusters of concepts within a concept map. The approach that we took to calculate concept densities is discussed in [2].

VOS

The construction of concept maps based on the association strengths of concepts was accomplished using a new method called VOS, which is an abbreviation for *visualization of similarities*. We now briefly introduce this method. A more elaborate discussion of VOS, including

[1]We note that a more detailed description of our research is available in [1].

an analysis of the relationship between VOS and multidimensional scaling, is provided in [3].

In this paper, the aim of VOS is to provide a two-dimensional space in which the concepts $1, \ldots, n$ are located in such a way that the distance between any pair of concepts i and j reflects their association strength a_{ij} as accurately as possible. Concepts that have a high association strength should be located close to each other, whereas concepts that have a low association strength should be located far from each other. The idea of VOS is to minimize a weighted sum of the squared Euclidean distances between all pairs of concepts. The higher the similarity between two concepts, the higher the weight of their squared distance in the summation. To avoid solutions in which all concepts are located at the same coordinates, the constraint is imposed that the sum of all distances must equal some positive constant. In mathematical notation, the objective function to be minimized in VOS is given by

$$E(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \sum_{i<j} a_{ij}\|\mathbf{x}_i - \mathbf{x}_j\|^2$$
(2)

where the vector $\mathbf{x}_i = (x_{i1}, x_{i2})$ denotes the coordinates of concept i in a two-dimensional space and $\|\cdot\|$ denotes the Euclidean norm. Minimization of the objective function is performed subject to the following constraint

$$\sum_{i<j} \|\mathbf{x}_i - \mathbf{x}_j\| = 1.$$
(3)

Note that the distances $\|\mathbf{x}_i - \mathbf{x}_j\|$ in the constraint are not squared. We numerically solved the optimization problem of minimizing (2) subject to (3) using the optimization toolbox of MATLAB.

Analysis

The thesaurus of the CI field that we used in our analysis contains 325 concepts. When constructing concept maps, concepts occurring in less than seven abstracts were excluded from further analysis. This was done because the amount of data on these concepts was considered too limited to calculate reli-
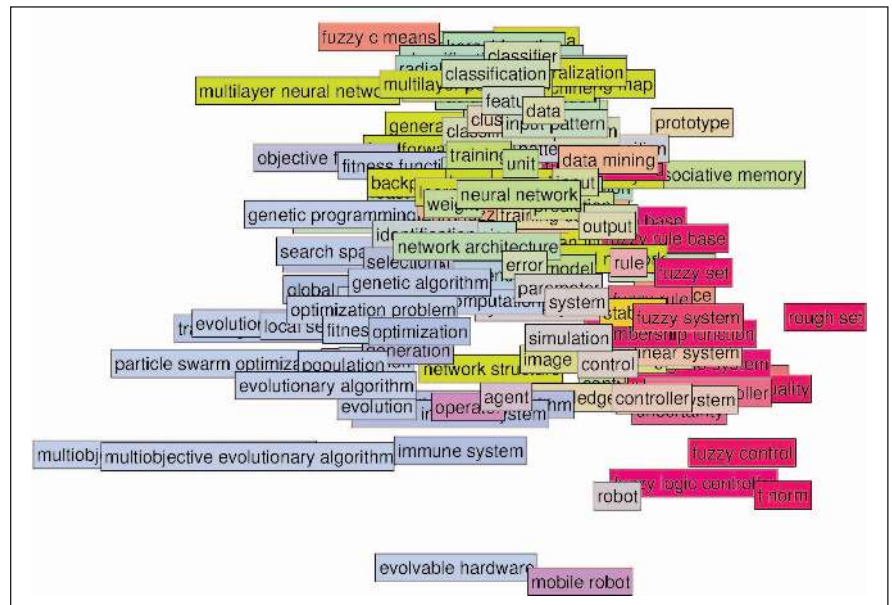

FIGURE 1 Concept map of the CI field in 2002.

able association strengths. This restriction resulted in a concept map of the CI field in 2002 that contains 134 concepts and a concept map of the CI field in 2006 that contains 172 concepts. The map for 2002 was constructed based on 1149 abstracts of the WCCI 2002, while the map for 2006 was constructed based on 1687 abstracts of the WCCI 2006.[2] The concept maps of the CI field in 2002 and 2006 are shown in Figures 1 and 3, respectively. The corresponding concept density maps are shown in Figures 2 and 4. We note that the concept maps in Figures 1 and 3 can be examined in much more detail using the concept map viewer that we have made available online.[3] The colors in the concept maps indicate the subconferences of the WCCI with which concepts are associated. Con-
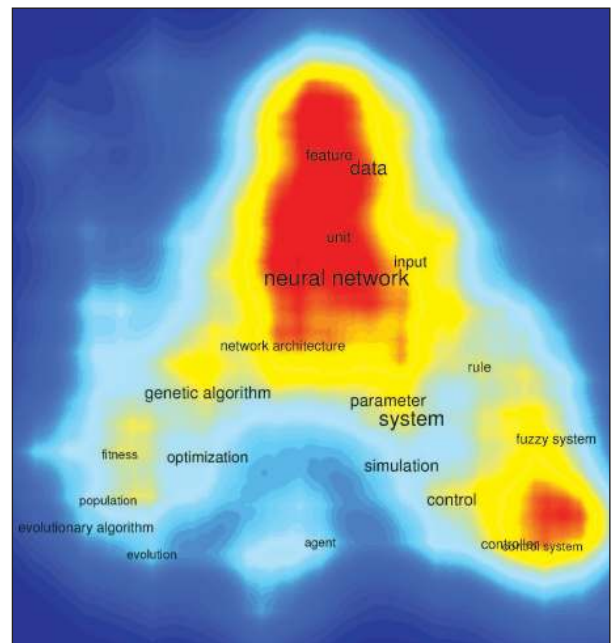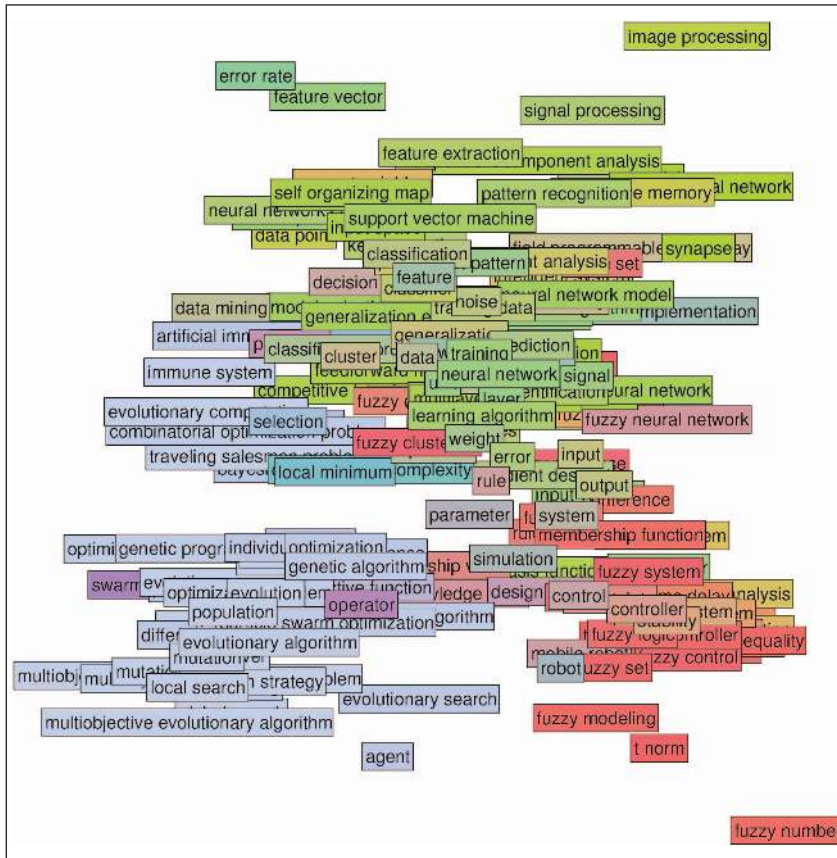

FIGURE 2 Concept density map of the CI field in 2002.

cepts occurring in abstracts of the International Joint Conference on Neural Networks (IJCNN) have green labels, while concepts occurring in abstracts of the IEEE International Conference on Fuzzy Systems (FUZZ-IEEE) and the IEEE Congress on Evolutionary Computation (CEC) have red and blue labels, respectively. When concepts occur in abstracts of more than one subconference of the WCCI, the colors of the subconferences

---

[2]At the time the abstracts were made available to us, there were a number of submissions to the WCCI 2006 for which a decision about acceptance or rejection had not yet been taken. We did not use the abstracts of these submissions in our analysis.

[3]See http://people.few.eur.nl/nvaneck/wcci2006/.

FIGURE 3 Concept map of the CI field in 2006.

FIGURE 4 Concept density map of the CI field in 2006.

cept map therefore seems to confirm the global picture that most researchers have of the structure of the CI field.

The concept density map in Figure 4 also provides insight into the relations between the three subfields of the CI field and into the characteristics of each of the subfields. The map indicates that the neural networks subfield and the fuzzy systems subfield are less separated from each other than from the evolutionary computation subfield. Compared to the other subfields, the evolutionary computation subfield therefore seems to have a rather independent position. The concept density map further reveals that the three subfields of the CI field have quite different characteristics. The neural networks subfield takes up a lot of space in the map, which suggests that this subfield covers a number of relatively diverse research topics. The fuzzy systems subfield, on the contrary, does not take up much space. The dark red color of the center of the fuzzy systems cluster indicates a large group of highly associated concepts. These concepts all turn out to be related to the topic of control, which appears to indicate that the fuzzy systems subfield is largely dominated by control research. The evolutionary computation subfield is similar to the neural networks subfield in the sense that it seems to cover several relatively diverse research topics. The coloring of the evolutionary computation cluster suggests that none of the topics dominates the others.

There is one other observation that we would like to make. The concept map in Figure 3 displays four red-colored concepts within the green colored neural networks cluster. These concepts are *fuzzy c-means*, *fuzzy classifier*, *fuzzy clustering*, and *rough set*. (The concepts may be difficult to see in Figure 3, but they can be easily found using the concept map viewer that is available online.) As indicated by their color, the four concepts are usually associated with the FUZZ-IEEE and, hence, with the fuzzy systems subfield. However, the concept map shows that the concepts are much more related to research

Structure of the Computational Intelligence Field

We analyze the structure of the CI field based on the concept map and the concept density map for 2006 (Figures 3 and 4). The concept density map clearly shows that concepts in the CI field are grouped into three clusters. Only a very limited number of concepts do not fit neatly into one of the clusters. Most researchers think of CI as a field that is divided into three subfields: neural networks, fuzzy systems, and evolutionary computation. It can be observed that the three clusters in the concept map correspond almost perfectly to these subfields. The con-

are mixed together in the appropriate proportion (for more details, see [1]). In the concept density maps, colors indicate the density of concepts. Dark blue indicates the lowest density and dark red indicates the highest density.

topics in the neural networks subfield than to the topic of control that dominates the fuzzy systems subfield. This could be interpreted as an indication that fuzzy systems is not a very homogeneous field. We further note that the concept map for 2002 (Figure 1) also reveals a strong relation between the concept *fuzzy c-means* and topics in the neural networks subfield.

### Evolution of the Computational Intelligence Field over the Last Years

We first analyze the differences in the frequency of occurrence of concepts in 2002 and 2006. In Table 1, the concepts are listed that have the largest relative increase in their frequency of occurrence between 2002 and 2006. Only concepts for which a statistical test indicated that the increase is significant at the 2.5% level are shown. For each concept, the number of abstracts in which the concept occurs in 2002 and 2006 is reported in the table. When interpreting the data in the table, one should of course keep in mind that for 2006 the total number of abstracts is almost 1.5 times as large as for 2002.

The data in Table 1 indicate a number of emerging areas within the CI field. Interestingly, most of these areas lie within the evolutionary computation subfield. The data reveal three emerging areas within this subfield. These areas are differential evolution, particle swarm optimization, and the application of evolutionary computation to multiobjective optimization problems. The latter area is indicated by the concept

NSGA-II. Another emerging area revealed by the data in Table 1 is support vector machines. Most abstracts containing the concept *support vector machine* belong to papers presented at the IJCNN, which shows that support vector machines research is usually seen as part of the neural networks subfield. Given the fairly large number of WCCI papers concerned with support vector machines, it is quite remarkable that the topic of support vector machines is not

covered in two recent textbooks on CI [4], [5]. Apparently there is no consensus within the CI community on the question whether support vector machines research belongs to the CI field at all. The data in Table 1 further seem to indicate that the interest of CI researchers in the topics of control design, principal component analysis, fuzzy clustering, and feature selection has increased considerably over the last years. We also note that, based on the frequencies of occurrence of concepts, no clear indications can be found of research topics in the CI field that receive much less attention in 2006 than they received in 2002.

We now compare the concept maps of the CI field in 2002 and 2006. The concept density maps in Figures 2 and 4 provide a quick overview of the main developments in the CI field over the last years. The maps clearly show that the evolutionary computation subfield has become more independent from the neural networks subfield. This seems to indicate that more and more research
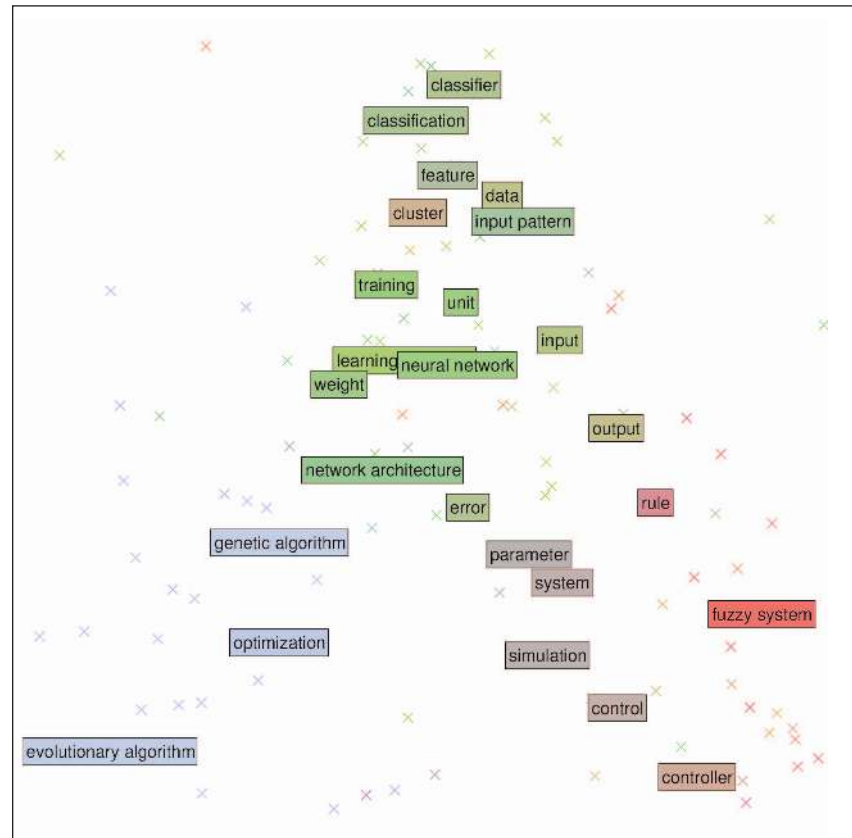
FIGURE 5 Simplified concept map of the CI field in 2002.

into evolutionary computation is concerned with evolutionary computation techniques on their own rather than with evolutionary computation techniques in combination with other CI techniques. The concept maps in Figures 1 and 3 can be used for a more detailed analysis of recent developments in the CI field. There are quite a lot of differences between the two maps. However, most differences have to do with concepts that have a low frequency of occurrence. The locations of these concepts in the maps may not be very accurate, since they are based on a rather limited amount of data. We therefore choose not to pay too much attention to these concepts. Instead, we focus on the 25 concepts with the highest frequency of occurrence in 2002. Simplified concept maps that display only these concepts are presented in Figures 5 and 6. We again observe that the intertwining of the evolutionary computation subfield with the neural networks subfield has decreased considerably over the last years. In the map for 2006, distances

between concepts from the evolutionary computation subfield on the one hand and concepts from the neural networks subfield on the other hand are larger than in the map for 2002. Especially the large increase in the distance between the concepts *genetic algorithm* and *network architecture* is quite remarkable. The locations of concepts like *cluster*, *data*, and *rule* also differ somewhat in the two maps. However, overall it can be concluded that, apart from the decreased association between evolutionary computation concepts and neural network concepts, the associations between the main concepts in the CI field have been fairly stable over the last years. This might be regarded as an indication that, despite its young age, the CI field is rapidly growing to maturity.

Conclusions

In this paper, we have visualized the structure and the evolution of the CI field. Based on our visualizations, we can draw a number of conclusions. The global picture that most researchers have

of the structure of the CI field seems correct. Our visualizations clearly indicate the well-known division of the field into the subfields of neural networks, fuzzy systems, and evolutionary computation. Nowadays the evolutionary computation subfield seems to have a rather independent position within the CI field. In our visualizations, the evolutionary computation subfield appears somewhat separated from the rest of the field. The neural networks subfield and the evolutionary computation subfield both seem to cover a number of relatively diverse research topics. The fuzzy systems subfield seems to be largely dominated by control research.

Over the last years, the structure of the CI field seems to have been quite stable, which might be regarded as an indication that the field is rapidly growing to maturity. The only indication that we have of a changing structure of the field is the more independent position of the evolutionary computation subfield in 2006 compared to 2002. We have also searched for emerging areas within the CI field. Based on a comparison of data from 2002 and 2006, we have identified four such areas. These areas are differential evolution, particle swarm optimization, the application of evolutionary computation to multiobjective optimization problems, and support vector machines. Interestingly, three out of the four areas lie within the evolutionary computation subfield, which suggests that this subfield has been particularly innovative over the last years. We further note that it is not completely clear whether the fourth area, support vector machines, belongs to the CI field at all.
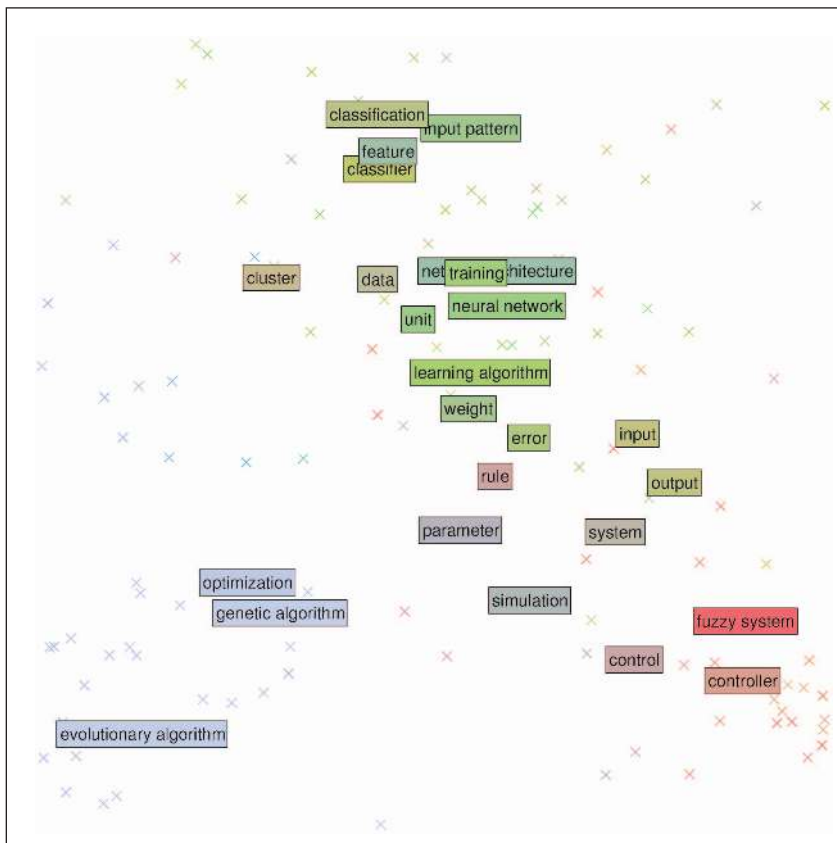
FIGURE 6 Simplified concept map of the CI field in 2006.

References
[1] N.J. van Eck, L. Waltman, J. van den Berg, and U. Kaymak, "Visualizing the WCCI 2006 knowledge domain," in *Proc. 2006 IEEE Int. Conf. Fuzzy Systems*, pp. 7862–7869, 2006.
[2] N.J. van Eck, F. Frasincar, and J. van den Berg, "Visualizing concept associations using concept density maps," in *Proc. 10th Int. Conf. Information Visualisation*, pp. 270–275, 2006.
[3] N.J. van Eck and L. Waltman, "VOS: A new method for visualizing similarities between objects," in *Proc. 30th Ann. Conf. German Classification Society*, 2006, accepted for publication.
[4] A.P. Engelbrecht, *Computational intelligence: an introduction*, John Wiley & Sons, 2003.
[5] A. Konar, *Computational intelligence: principles, techniques and applications*, Springer, 2005.