J Inf Process Syst, Vol.13, No.3, pp.435~447, June 2017
https://doi.org/10.3745/JIPS.02.0058

ISSN 1976-913X (Print)
ISSN 2092-805X (Electronic)

INVITED PAPER

JOURNAL OF INFORMATION PROCESSING SYSTEMS JIPS

Granular Bidirectional and Multidirectional Associative Memories: Towards a Collaborative Buildup of Granular Mappings

Witold Pedrycz*

Abstract

Associative and bidirectional associative memories are examples of associative structures studied intensively in the literature. The underlying idea is to realize associative mapping so that the recall processes (one-directional and bidirectional ones) are realized with minimal recall errors. Associative and fuzzy associative memories have been studied in numerous areas yielding efficient applications for image recall and enhancements and fuzzy controllers, which can be regarded as one-directional associative memories. In this study, we revisit and augment the concept of associative memories by offering some new design insights where the corresponding mappings are realized on the basis of a related collection of landmarks (prototypes) over which an associative mapping becomes spanned. In light of the bidirectional character of mappings, we have developed an augmentation of the existing fuzzy clustering (fuzzy c-means, FCM) in the form of a so-called collaborative fuzzy clustering. Here, an interaction in the formation of prototypes is optimized so that the bidirectional recall errors can be minimized. Furthermore, we generalized the mapping into its granular version in which numeric prototypes that are formed through the clustering process are made granular so that the quality of the recall can be quantified. We propose several scenarios in which the allocation of information granularity is aimed at the optimization of the characteristics of recalled results (information granules) that are quantified in terms of coverage and specificity. We also introduce various architectural augmentations of the associative structures.

Keywords

Allocation of Information Granularity and Optimization, Bidirectional Associative Memory, Collaborative Clustering, Granular Computing, Multi-directional Associative Memory, Prototypes

1. Introduction

Associative memories have been intensively studied for several decades being driven by several fundamental and applied research. The fundamentals origin from the interest in intriguing ways for understanding how memories organize, store, and retrieve data. There are some biologically-inclined studies that also play a visible role [1-3]. The crux of the recall in associative memories is to produce (recall) an item associated with some available chunk of data being incomplete or noisy.

The term "association" plays a pivotal role in human endeavors and goes back to the ideas of Hebbian

Manuscript received April 17, 2017; accepted May 15, 2017.
Corresponding Author: Witold Pedrycz (wpedrycz@ualberta.ca)
* Dept. of Electrical & Computer Engineering University of Alberta Edmonton T6R 2V4 AB Canada (wpedrycz@ualberta.ca); Dept. of Electrical and Computer Engineering Faculty of Engineering, King Abdulaziz University Jeddah, Saudi Arabia; Systems Research Institute, Polish Academy of Sciences Warsaw, Poland

(correlation) learning and so-called correlation associative memories. The studies on associative memories reported in the existing literature revolve along concepts such as stability, recall properties (error), capacity, and robustness. Central to associative memories are the issues of knowledge representation, design practices, performance analysis, and optimization.

The advances in fuzzy sets and neurocomputing have resulted in the emergence of so-called fuzzy associative memories (refer to [4-6], which highlight the main representative architectures of fuzzy associative memories). The pertinent design practices are discussed in [7-14], which all address the issues of stability, recall quality, and neural network-based realizations. The development of associative memories comes with a variety of frameworks, including logic-oriented [15] and quantum logic environments [16,17]. Cellular-based automata approach was studied in [18], graph-based approaches were presented in [19], and hierarchical structures were studied in [20]. The wealth of applications is deep and includes tackling parameter estimation and control problems [21,22], decision-support systems [23], and formal concept analysis [24].

This study contributes to the conceptual, architectural, and design–oriented enhancements of associative memories, and as such comes with several well-defined objectives that are implied by the established conceptual and application-driven goals listed below.

(i) The construction of associations focused on the key structural dependencies among the data by spanning the associative memory (associative mapping) over the representatives of the data. We are proposing an algorithmic setting realized in the form of collaborative fuzzy clustering.

(ii) The quantification of the quality of associative memories by bringing the ideas of information granules. Subsequently, we were able to deliver an efficient way of quality description of recall through the coverage and specificities that characterize the quality of recall results.

(iii) The extension of the notion of bidirectionality to multi-directionality and a proposal of architectural extensions of bidirectional memories to network topologies of associative memories.

All of these objectives exhibit a significant deal of originality by opening some new avenues of research and applications.

The paper is structured in a way that we unveil the main concepts and ensuing algorithmic design practices. In Section 2, we start with by briefly highlighting the essence of associative mapping by characterizing it as an optimization problem. We advocate for its structure by showing that the mapping spans a collection of representatives of data for which associations are to be built. We identify some conceptual advantages and deliver an efficient way of building memories. Section 3 is focused on the algorithm behind the buildup of the structural core of memories, which creates a collaborative version of clustering followed by the optimization problem arising in conjunction of the optimal recall (Section 4). An allocation of information granularity providing an efficient way of describing the quality of the associative memories is formulated in Section 5 and then supplied with solutions (Section 6). Multi-directional associative memories are discussed in Section 7, and we present out conclusions in Section 8.

2. Bidirectional Associative Memories: Building Mappings Spanned over Data Representatives

In a nutshell, a bidirectional associative memory can be regarded as a pair of mappings f: $\boldsymbol{X} \rightarrow \boldsymbol{Y}$ and g: $\boldsymbol{Y} \rightarrow \boldsymbol{X}$ where, $\boldsymbol{X}$ and $\boldsymbol{Y}$ are subsets of associated items (data) positioned in the multidimensional spaces of real numbers $\boldsymbol{R}^n$ and $\boldsymbol{R}^m$, respectively. The mappings are formed in such a way that they produce minimal recall errors for any finite number of pairs of items stored in the memory. More specifically, f and g are constructed in such a manner that for all pairs of stored items (associations) ($\boldsymbol{x}_k$, $\boldsymbol{y}_k$) $\boldsymbol{x}_k \in \boldsymbol{R}^n, \boldsymbol{y}_k \in \boldsymbol{R}^m$, k=1,2,…, N, the following recall conditions are (approximately) satisfied, $f(\boldsymbol{x}_k) \approx \boldsymbol{y}_k$ and $g(\boldsymbol{y}_k) \approx \boldsymbol{x}_k$, k=1,2,…, N. It is not possible to anticipate that the associative recall could be ideal. Some recall errors occur, and as a result, the ultimate objective of the overall design process is to develop the structure of the mappings (f and g) and to optimize their parameters in such a way that a certain performance index V (commonly shown as the sum of the squared errors $||.||^2$) becomes minimized as follows:

$$V = \sum_{k=1}^{N} || \boldsymbol{x}_k - \hat{\boldsymbol{x}}_k ||^2 + \sum_{k=1}^{N} || \boldsymbol{y}_k - \hat{\boldsymbol{y}}_k ||^2 \quad (1)$$

where, $\hat{\boldsymbol{x}}_k, \hat{\boldsymbol{y}}_k$ are the recall results of the associated item (data), namely $f(\boldsymbol{x}_k) = \hat{\boldsymbol{y}}_k$ and $g(\boldsymbol{y}_k) = \hat{\boldsymbol{x}}_k$. Note that we assessed the recall quality in both directions, viz. for the items in $\boldsymbol{X}$ and $\boldsymbol{Y}$.

The aim of the optimization, as noted earlier, is to minimize V by selecting the mappings from some classes of admissible mappings $\boldsymbol{F}$ and $\boldsymbol{G}$ and estimating the values of the parameters of f and g, say $f(\boldsymbol{x},\boldsymbol{a})$ and $g(\boldsymbol{y}, \boldsymbol{b})$. This results in the following result of optimization:

$$(f_{\text{opt}}, g_{\text{opt}}, \boldsymbol{a}_{\text{opt}}, \boldsymbol{b}_{\text{opt}}) = \arg Max_{f \in \boldsymbol{F}, g \in \boldsymbol{G}; \boldsymbol{a}, \boldsymbol{b}} V \quad (2)$$

What we are presenting here constitutes a radical shift from the main line of study in the sense that we are advocating that the essence and generality of associative mappings shouldn't reflect numerous amounts of data but instead to build upon a collection of representatives of the association pairs. By focusing on the buildup of the associative memory on these landmarks, two motivating factors are worth pointing out, which are listed below.

(i) We can avoid an immediate and direct reliance on noisy and distorted data. The prototypes are a kind of summarization of the data, and in this way, they capture the essence of the data and tend to ignore details that are most likely a result of noise manifestation in the data.

(ii) The structure of the mapping focuses on the prototypes and the choice of the prototypes helps set up a certain tradeoff between detailed recall and noise immunities.

Once the prototypes (representatives) have been formed, the details of the mapping are developed in terms of the detailed structure of the associative linkages, especially the parameters of the mappings and the way in which some interpolation/approximation capabilities are determined.

From the design perspective, the problem naturally splits into two fundamental phases, as described below.

(i) Building representatives (landmarks) of associative mappings and their refinements (optimization). This includes the realization of a certain form of the associative mapping spanned over a small number of representatives – information granules.

(ii) The characterization of the quality (performance) of the associative memories. We demonstrate in this section that the memory's quality can be efficiently described by means of information granules. By augmenting the existing numeric constructs and bringing a facet of information granularity (as a consequence yielding so-called granular associative memories), we deliver a quantitative view at the performance of the memory.

Next, we concisely present the underlying essence of our proposed approach. The landmarks that form the backbone of the associative mapping are developed through clustering (or fuzzy clustering) the data in ***X*** and ***Y***, as shown in Fig. 1.

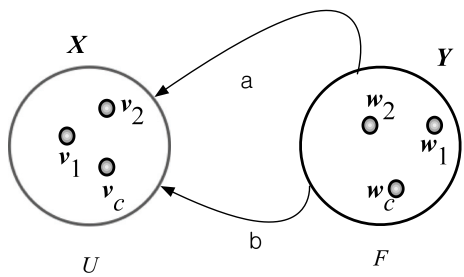

Fig. 1. A concept of bidirectional recall realized with mappings *f* and *g* spanned over a collection of *c* prototypes; *U* and *F* stand for the partition matrices; α, β strengths of the collaboration.

The prototypes that are formed because of clustering serve as the entities over which the mapping is formed. Clustering is a common way of determining the representatives of the overall dataset, where the obtained prototypes serve as points for which the associations are developed. In contrast to the commonly studied way of forming clusters, due to the clustering that needs to be completed in concert for data in ***X*** and ***Y*** at the same time, the generic scheme of clustering needs to be prudently revisited so that the prototypes in ***X*** and ***Y*** are determined (and optimized) together. Doing so calls for clustering that is completed in a collaborative manner. Once the prototypes have been constructed they are used to finalize the mapping. In this study, we provide a rule-based architecture of the associative mapping. The form of the mapping is realized on the basis of a collection of reference points (prototypes) $\boldsymbol{v}_1$, $\boldsymbol{v}_2,\ldots,\boldsymbol{v}_c$ that are defined in ***X***, and the corresponding reference points (prototypes) $\boldsymbol{w}_1$, $\boldsymbol{w}_2,\ldots$, $\boldsymbol{w}_c$ are defined in ***Y***.

3. Collaborative Clustering and its Underlying Algorithm

In the context of the requirements imposed on the formation of representatives (prototypes) of the data ***X*** and ***Y***, we are introducing a generalized version of fuzzy clustering that can be regarded as a generalized version of fuzzy c-means (FCM). The use of FCM in the context of this study is beneficial because of the common usage of this clustering technique, especially in the realm of fuzzy set constructs (classifiers, predictors, control, etc.). To fully align the clustering procedure with the resulting associative mapping, we referred to the notation provided in Fig. 1.

The objective function that guides the clustering process is described as follows:

$$Q=\frac{1}{2}\sum_{i=1}^{c}\sum_{k=1}^{N}u_{ik}^{2}d_{ik}^{2}+\frac{1}{2}\alpha\sum_{i=1}^{c}\sum_{k=1}^{N}(u_{ik}-f_{ik})^{2}d_{ik}^{2} \tag{3}$$

This objective function is concerned with the clustering of data ***X***. Here, $U=[u_{ik}]$ and $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, .., $\boldsymbol{v}_c$ are the partition matrixes and prototypes, respectively. The expression d_{ik} stands for the Euclidean distance (or its weighted version) between $\boldsymbol{x}_k$ and the prototype $\boldsymbol{v}_i$. The prototypes positioned in $\boldsymbol{Y}$ are $\boldsymbol{w}_1$, $\boldsymbol{w}_2$,…, $\boldsymbol{w}_c$; whereas, the partition matrix is $F=[f_{ik}]$, i=1,2,.., c; k=1,2,.., N. The structure of the objective function (3) consists of two components. The first one is related to the structure being revealed in $\boldsymbol{X}$. The second one shows a level of agreement between the structures discovered simultaneously in $\boldsymbol{X}$ and $\boldsymbol{Y}$.

To proceed with optimization, we formed an augmented objective function Q_1 by including a Lagrange multiplier to accommodate the constraint of the partition matrix:

$$Q_1 = \frac{1}{2}\sum_{i=1}^{c} u_{ik}^2 d_{ik}^2 + \frac{1}{2}\alpha\sum_{i=1}^{c}(u_{ik}-f_{ik})^2 d_{ik}^2 + \lambda(\sum_{i=1}^{c} u_{ik} - 1) \tag{4}$$

The derivatives of Q_1 computed with respect to the entries of the partition matrix and the prototypes are set to zero, which are necessary conditions of the minimum of Q_1:

$$\frac{dQ_1}{du_{st}} = u_{st}d_{st}^2 + \alpha(u_{st}-f_{st})d_{st}^2 + \lambda = 0 \tag{5}$$

Then:

$$u_{st} = \frac{\alpha f_{st} d_{st}^2 + \lambda}{d_{st}^2 + \alpha d_{st}^2} \tag{6}$$

As the sum of membership grades is equal to 1, $\sum_{s=1}^{c} u_{st}$ =1 we obtained:

$$\sum_{s=1}^{c} \frac{\alpha f_{st} d_{st}^2 + \lambda}{d_{st}^2 + \alpha d_{st}^2} = 1 \tag{7}$$

and after some straightforward algebra one obtains:

$$\lambda = \frac{1}{\sum_{s=1}^{c}\frac{1}{d_{st}^2}} \tag{8}$$

Plugging (8) into (6) and the membership values of the partition matrix are finally obtained as:

$$u_{ik} = \frac{\alpha}{1+\alpha} f_{ik} + \frac{1}{1+\alpha}\frac{1}{\sum_{j=1}^{c}\frac{d_{ik}^2}{d_{jk}^2}} \tag{9}$$

i=1, 2, …,c, k=1, 2, …, N. By zeroing the derivative dV/dv_{ij}=0, i=1, 2,…, c, j=1, 2,…, n, the optimal prototypes are computed as follows:

$$\boldsymbol{v}_i = \frac{\sum_{k=1}^{N}[u_{ik}^2 + \alpha(u_{ik}-f_{ik})^2]\boldsymbol{x}_k}{\sum_{k=1}^{N}[u_{ik}^2 + \alpha(u_{ik}-f_{ik})^2]} \tag{10}$$

i=1,2,…, c.

Let us take a closer look at (9). If $\alpha = 0$, we arrive at the partition matrix that is encountered in the standard FCM (viz. no collaboration present, α=0), namely:

$$u_{st} = \frac{1}{\sum_{i=1}^{c} \frac{d_{st}^2}{d_{it}^2}} \tag{11}$$

If α=1, the partition matrix is equally impacted by the structure produced by $\boldsymbol{X}$ without any collaboration and the partition matrix F reflecting a structure discovered in $\boldsymbol{Y}$. If α tends to infinity, u_{ik} is equal to f_{ik}, which means that the structure in $\boldsymbol{X}$ is completely imposed by the findings present in $\boldsymbol{Y}$ in terms of the partition F.

For the collaborative clustering for the second data of $\boldsymbol{Y}$, the formulation of the problem is presented in the same way. We consider the objective function of:

$$Q = \frac{1}{2}\sum_{i=1}^{c}\sum_{k=1}^{N} f_{ik}^2 d_{ik}^2 + \frac{1}{2}\beta\sum_{i=1}^{c}\sum_{k=1}^{N}(f_{ik}-u_{ik})^2 d_{ik}^2 \tag{12}$$

Here, the distance d_{ik} denotes the Euclidean distance computed for the data $\boldsymbol{y}_k$ and the prototype $\boldsymbol{w}_k$.

Proceeding with the same optimization as before, the obtained detailed formulas read as follows:

Partition matrix F:

$$f_{ik} = \frac{\beta}{1+\beta} f_{ik} + \frac{1}{1+\beta} \frac{1}{\sum_{j=1}^{c} \frac{d_{ik}^2}{d_{jk}^2}} \tag{13}$$

in the first term, it should be u_{ik} instead of f_{ik} prototypes:

$$\boldsymbol{w}_i = \frac{\sum_{k=1}^{N} [f_{ik}^2 + \beta(f_{ik}-u_{ik})^2]\boldsymbol{y}_k}{\sum_{k=1}^{N} [f_{ik}^2 + \beta(f_{ik}-u_{ik})^2]} \tag{14}$$

We observed that the results of collaborative clustering (partition matrices and prototypes) are impacted by the intensity of collaboration (quantified in regards to the non-zero collaboration coefficients α and β). Obviously, the collaboration need not to be symmetric, and, hence, the values α and β can be independently set up. Choosing the values of these coefficients impacts the results by moving the position of the obtained prototypes. This flexibility can be utilized to further optimize the performance of the associative mappings.

4. Optimization Mechanism in Collaborative Clustering

The prototypes $\boldsymbol{v}_i$ and $\boldsymbol{w}_i$ exhibit a direct impact on the performance of the associative recall. The performance is quantified by summing up the distances between the original items and their associative recalls. As we are concerned with bidirectional recall, the sum of errors V is composed of the two parts that quantify the difference between the original data and their recalled versions. The recall mappings that span the developed prototypes are expressed as follows:

Recall completed from $\boldsymbol{X}$ to $\boldsymbol{Y}$:

$$\hat{\boldsymbol{y}}_k = \sum_{i=1}^{c} A_i(\boldsymbol{x}_k)\boldsymbol{w}_i \tag{15}$$

Recall completed from $\boldsymbol{Y}$ and $\boldsymbol{X}$:

$$\hat{\boldsymbol{x}}_k = \sum_{i=1}^{c} B_i(\boldsymbol{y}_k)\boldsymbol{v}_i \tag{16}$$

where, A_i and B_i stand for the activation levels (membership grades) implied by the data in $\boldsymbol{X}$ and $\boldsymbol{Y}$. More specifically:

$$A_i(\boldsymbol{x}_k) = \frac{1}{\sum_{j=1}^{c} \frac{\|\boldsymbol{x}_k - \boldsymbol{v}_i\|^2}{\|\boldsymbol{x}_k - \boldsymbol{v}_j\|^2}} \quad B_i(\boldsymbol{y}_k) = \frac{1}{\sum_{j=1}^{c} \frac{\|\boldsymbol{y}_k - \boldsymbol{w}_i\|^2}{\|\boldsymbol{y}_k - \boldsymbol{w}_j\|^2}} \tag{17}$$

Formally, for the fixed number of prototypes (c) the optimization problem is expressed as the minimization of the collaboration coefficients α and β, as:

$$(\alpha_{\text{opt}}, \beta_{\text{opt}}) = \arg Min_{\alpha,\beta \geq 0} V \tag{18}$$

The optimal values of the collaboration coefficients can be determined by using a population-based optimization (genetic algorithm, particle swarm, differential evolution, etc.), which is a sound alternative to the commonly used gradient-based techniques that are not feasible to use due to the non-explicit nature of the gradient of V with respect to the optimized parameters. As there are only two arguments involved, a simple sweeping across the values of α and β is also a viable optimization alternative.

5. Augmentation of Memories Through the Allocation of Information Granularity

As expected, there are no ideal bidirectional associative memories in the sense that the recall error is not equal to zero no matter how complex the associative memory is. To deal with this problem, we generalized the construct to make it granular, which resulted in a 'granular' associative memory. The concepts of information granularity and information granules are fundamental in this setting. Information granularity [25] serves as an important design asset. This augmentation is accomplished by making the prototypes in the form of information granules that span the original numeric counterparts, such as $\boldsymbol{V}_i = G(\boldsymbol{v}_i)$ and $\boldsymbol{W}_i = G(\boldsymbol{w}_i)$, where, G denotes a mechanism of forming an information granule applied to $\boldsymbol{v}_i$ and $\boldsymbol{w}_i$. In the simplest version, information granules can be presented in the form of intervals (sets). This yields the recall results that occur in the form of information granules. We have recall completed from $\boldsymbol{X}$ to $\boldsymbol{Y}$:

$$Y = \sum_{\substack{i=1 \\ \oplus}}^{c} [A_i(\boldsymbol{x}) \otimes G(\boldsymbol{v}_i)] = \sum_{\substack{i=1 \\ \oplus}}^{c} [A_i(\boldsymbol{x}) \otimes \boldsymbol{V}_i] \tag{19}$$

recall completed from $\boldsymbol{Y}$ to $\boldsymbol{X}$:

$$Y = \sum_{\substack{i=1 \\ \oplus}}^{c} [B_i(\boldsymbol{y}) \otimes G(\boldsymbol{w}_i)] = \sum_{\substack{i=1 \\ \oplus}}^{c} [B_i(\boldsymbol{y}) \otimes \boldsymbol{W}_i] \tag{20}$$

where, X and Y are information granules of the recalled item, which are formed on the basis of the already obtained numerical recall results. The symbols of addition and multiplication in circles stress that the arguments are information granules, not numbers. The operations of addition and multiplication are implemented depending upon the formalism involved. In the case of intervals, there are the operations of interval addition and multiplication. In case of fuzzy sets, we were concerned with the algebraic operations governed by the extension principle. Note that the equations presented above for the recalled items are information granules due to the granular form of the prototypes. The level of information granularity (specificity) of $\boldsymbol{W}_i$ and $\boldsymbol{V}_i$ is implied by the introduction a positive level of information granularity ε, which makes the prototypes granular. The quality of the granular recall is quantified by counting how many times the result of the recall X_k (and Y_k) includes (covers) the corresponding data $\boldsymbol{y}_k$ and $\boldsymbol{x}_k$ and how specific the recall results become. To explicitly articulate the level of information granularity allocated to the numeric prototypes, we used $\boldsymbol{W}_i(\varepsilon_i)$, with ε_i representing the level of information granularity of $\boldsymbol{W}_i$. This entails that the resulting recall is an information granule of the corresponding level of information granularity δ. To stress the granular character of the recall results, we rewrote (19) as $Y(\delta) = \sum_{\substack{i=1 \\ \oplus}}^{c} [A_i(\boldsymbol{x}) \otimes G(\boldsymbol{v}_i, \varepsilon_i)] = \sum_{\substack{i=1 \\ \oplus}}^{c} [A_i(\boldsymbol{x}) \otimes \boldsymbol{V}_i(\varepsilon_i)]$.

Proceeding with the detailed formulas, we have the following characterization of information granules: average coverage for recalled items in $\boldsymbol{Y}$:

$$\mathrm{cov}(\boldsymbol{Y}) = \frac{1}{N} \sum_{k=1}^{N} \mathrm{cov}(\boldsymbol{y}_k, Y_k) \tag{21}$$

In the case of set information granules, $\mathrm{cov}(\boldsymbol{y}_k, Y_k)$ is a binary predicate returning 1 if $\boldsymbol{y}_k$ is included in Y_k and is 0 otherwise.

The average specificity of recalled items is in $\boldsymbol{Y}$. The specificity is the measure that quantifies how detailed a certain information granule is. For a single element information granule, the specificity attains 1 and declines when the size of the information granule increases (higher values of ε_i). The specificity of the recall being completed in $\boldsymbol{Y}$ is determined as follows:

$$sp(\boldsymbol{Y}) = \frac{1}{N} \sum_{k=1}^{N} sp(Y_k) \tag{22}$$

Computing the quality of information granules for the recall completed in ***X*** is carried out in the same manner.

The quality of the recall is implied by the high average coverage and by the high average value of the specificity measure, and these two measures are in conflict. A simple alternative is to take a product of these two as:

$$T = \text{cov}(\boldsymbol{Y})*\text{sp}(\boldsymbol{Y}) + \text{cov}(\boldsymbol{X})*\text{sp}(\boldsymbol{X}) \tag{23}$$

The values of T depend upon the values of the level of information granularity. The higher values of δ (implied by higher values of ε_i) make the coverage higher, and at the same time, the specificity is reduced. The plotting of the average specificity-average coverage is shown in Fig. 2, which provides a visualization of the relationships between these two performance indicators and their conflicting nature. Furthermore, the plot completed in the coverage-specificity coordinates is beneficial to identifying the level of information granularity where the two criteria are still satisfied to high extent building in this way a sound compromise.

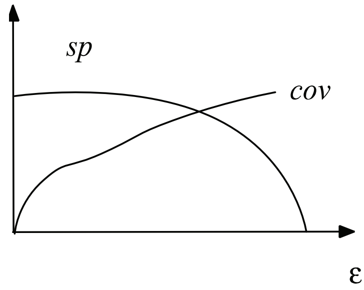

Fig. 2. Characteristics of coverage-specificity implied by various levels of information granularity ε.

6. Two Ways of Forming Granular Associative Memories

We are proposing to fundamentally different ways of forming information granules on the basis of the numeric prototypes. In short, the granule has to be reflective of the characteristics of the data ***X*** and ***Y***.

Uniform allocation of information granularity

The prototypes are made granular by using a certain level of information granularity ε and forming granular prototypes around their numeric counterparts. The granular prototypes are built as a geometric construct governed by the equation:

$$\mathbf{V}_i(\varepsilon) = \{\boldsymbol{x} \mid ||\boldsymbol{x}-\boldsymbol{v}_i||^2 \le n\varepsilon^2\} \tag{24}$$

It is noticeable that assuming the weighted Euclidean distance, the geometry of information granules is a circular shape. All prototypes come with the same level of information granularity ε. The choice of a suitable value of ε can be guided by analyzing the quality of granular results of recall by analyzing the coverage of the specificity plot, as shown in Fig. 2, and by selecting ε for which one achieves high values of coverage without substantial reductions in the values of specificity.

Formation of information granules on the basis of the local data characteristics

In this method, the granular prototype is formed by invoking the principle of justifiable granularity [25]. This principle supports the development of justifiable information granules, viz. the granules that are supported by experimental evidence while being semantically sound, viz. specific enough. The optimization of these two already discussed criteria gives rise to the information granule.

The advantage of this method comes with the explicit reliance on the local characteristics of the data positioned around the prototypes. This makes the granular prototypes reflective of the nature of the data and of different levels of information granularity, where we have $\boldsymbol{V}_i(\varepsilon_i)$ and $\boldsymbol{W}_i(\varepsilon_i)$. A certain shortcoming is the inherent computing overhead.

7. Multi-Directional Associative Memories: Selected Categories of Topologies

The associative memory can assume several key topologies that reflect the nature of the problem and in a way in which associations are established between data sources, as shown in Fig. 3.

One to one associative memory

Both one–directional and bidirectional structures are envisioned. Two sources of data $\boldsymbol{X}$ and $\mathbf{Y}$ are considered, and they have already been discussed in this study. The optimization problem is focused on the building landmarks of the mapped minimizing recall error. This situation is typical in system modeling where a one-way of association is typical for any model. More specifically, what has been presented above is a Takagi-Sugeno fuzzy rule-based model. To highlight the essence of the developed construct, note that $A_i(\boldsymbol{x})$ is a degree of matching (activation) of the i-th reference (landmark) point $\boldsymbol{v}_i$. The conclusion part of the i-th rule is a constant function that is formed as the prototype $\boldsymbol{w}_i$. The number of rules is equal to the number of clusters. If we view $\boldsymbol{X}$ as the space of independent variables, the recall completed in the direction from $\boldsymbol{X}$ to $\boldsymbol{Y}$ corresponds to the problem of the determination of the output. If the values of $\boldsymbol{y}$ that are located in $\boldsymbol{Y}$ are given and treated as a vector of effects, then the recall in $\boldsymbol{X}$ corresponds to the resulting vector of causes.

Many to one and one to many associative memories

In this case, we encountered a multiplicity of data sources $\boldsymbol{X}_1$, $\boldsymbol{X}_2$,…, X_r and any $\boldsymbol{Y}_1$, $\boldsymbol{Y}_2$,…, $\boldsymbol{Y}_p$. One typical alternative is to have a number of data sources $\boldsymbol{X}_1$,…, $\boldsymbol{X}_r$ interacting with a single data source $\boldsymbol{Y}$. In terms of applications, the output located in the output space $\boldsymbol{Y}$ is a result of considering many factors. Associations are presented and separately developed in the spaces $\boldsymbol{X}_1$,…, $\boldsymbol{X}_r$ and are mapped on $\boldsymbol{Y}$. This option can be linked with systems in which a certain input space comes in a collection of subspaces that are available locally and the output is observed in $\boldsymbol{Y}$. The other option is visualized in Fig. 3(b), where a single data space $\boldsymbol{X}$ implies the outputs that are localized in $\boldsymbol{Y}$ and $\boldsymbol{Z}$.

Multi-level structure of associations

In this case, we envisioned a topology where associations among spaces are cascaded. An illustrative example is shown in Fig. 3(c), where associations are built between ***X*** and ***Z*** and ***Z*** and ***Y***, and the number of intermediate levels could be higher. It is worth noting that such situations occur when we do not have any direct linkage between ***X*** and ***Y*** and they can be formed indirectly with the aid of the associations in ***X*** and ***Z*** and ***Z*** and ***Y***. In other words, this option is pursued when ***X*** and ***Y*** are not present together and the structure could be revealed by invoking the intermediate data source ***Z*** and forming the related associations.

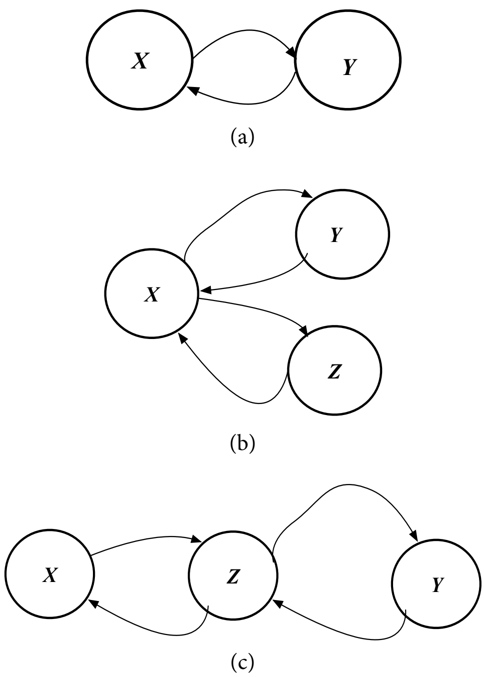

Fig. 3. Examples of multi-directional associations.

As there are a larger number of data spaces and ensuing associations to be established, there are a larger number of collaboration coefficients to deal with when clustering. Here, a viable optimization framework is to resort to genetic algorithms (GAs), particle swarm optimizations (PSOs), or other techniques located in this category of population-based algorithms.

8. Conclusions

In this study, we have provided a new direction in the design of bidirectional associative memories. The proposed development environment comes with several outstanding features. First, the construction of the best mapping is formed with the aid of landmarks (prototypes) whose position in the data spaces captures the key properties of the data. The underlying clustering technique involves the mechanism of collaboration, which helps distribute the prototypes in such a way that the bidirectionality (or multi-directionality) of the recall is fully reflected. Second, the quality of the nonideal associative recall is quantified through information granules, and we showed that a level of information granularity can be optimized by the criteria of data coverage and specificity. Third, the concept of multi-directional recall was introduced, which is a result of a generalized multi-source of networks of associations that extend far beyond the classic architectures of the existing associative memories.

Future research studies can naturally revolve around further investigations of the ideas introduced here, especially at the level of algorithmic details. Different formalisms of information granules beyond the intervals that we have addressed in this study can be explored. There are also further applied studies that could be carried out, in particular in data analytics.

Acknowledgement

Support from the Canada Research Chair (CRC) Program and the Natural Sciences and Engineering Research Council (NSERC) are gratefully acknowledged.

References

[1] E. K. Bridger, A. L. Kursawe, R. Bader, R. Tibon, N. Gronau, D. A. Levy, and A. Mecklinger, "Age effects on associative memory for novel picture pairings," *Brain Research*, vol. 1664, pp. 102–115, 2017.

[2] G. Palm, "On associative memory," *Biological Cybernetics*, vol. 36, no. 1, pp. 19–31, 1980.

[3] G. Palm, "Neural associative memories and sparse coding," *Neural Networks*, vol. 37, pp. 165–171, 2013.

[4] R. Belohlavek, "Fuzzy logical bidirectional associative memory," *Information Sciences*, vol. 128, no. 1, pp. 91–103, 2000.

[5] B. Kosko, "Bidirectional associative memories," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 18, no. 1, pp. 49–60, 1988.

[6] M. E. Valle and P. Sussner, "Storage and recall capabilities of fuzzy morphological associative memories with adjunction-based learning," *Neural Networks*, vol. 24, no. 1, pp. 75–90, 2011.

[7] Q. Cheng and Z. T. Fan, "The stability problem for fuzzy bidirectional associative memories," *Fuzzy Sets and Systems*, vol. 132, no. 1, pp. 83–90, 2002.

[8] E. Esmi, P. Sussner, and S. Sandri, "Tunable equivalence fuzzy associative memories," *Fuzzy Sets and Systems*, vol. 292, pp. 242–260, 2016.

[9] W. Feng, S. X. Yang, H. Wu, "Further results on robust stability of bidirectional associative memory neural networks with norm-bounded uncertainties," *Neurocomputing*, vol. 148, pp. 535–543, 2015.

[10] S. C. Hana, Y. D. Gu, and H. X. Li, "An application of incline matrices in dynamic analysis of generalized fuzzy bidirectional associative memories," *Fuzzy Sets and Systems*, vol. 158, no. 12, pp. 1340–1347, 2007.

[11] O. Qadir, J. Liu, G. Tempesti, J. Timmis, and A. Tyrrell, "From bidirectional associative memory to a noise-tolerant, robust protein processor associative memory," *Artificial Intelligence*, vol. 175, no. 2, pp. 673–693, 2011.

[12] F. Shen, Q. Ouyang, W. Kasai, and O. Hasegawa, "A general associative memory based on self-organizing incremental neural network," *Neurocomputing*, vol. 104, pp. 57–71, 2013.

[13] M. E. Valle, "Permutation-based finite implicative fuzzy associative memories," *Information Sciences*, vol. 180, no. 21, pp. 4136–4152, 2010.

[14] C. Zhong, W. Pedrycz, Z. Li, D. Wang, and L. Li, "Fuzzy associative memories: a design through fuzzy clustering," *Neurocomputing*, vol. 173, pp. 1154–1162, 2016.

[15] J. H. Qiang, W. X. Xin, and T. J. Feng, "Improved dynamic subjective logic model with evidence driven," *Journal of Information Processing Systems*, vol. 11, no. 4, pp. 630–642, 2015.

[16] G. G. Rigatos and S. G. Tzafestas, "Quantum learning for neural associative memories," *Fuzzy Sets and Systems*, vol. 157, no. 3, pp. 1797–1813, 2006.

[17] D. Ventura and T. Martinez, "Quantum associative memory," *Information Sciences*, vol. 124, no. 1, pp. 273–296, 2000.

[18] P. Maji and P. P. Chaudhuri, "Non-uniform cellular automata based associative memory: evolutionary design and basins of attraction," *Information Sciences*, vol. 178, no. 10, pp. 2315–2336, 2008.

[19] A. D. Bouchain and G. Palm, "Neural coding in graphs of bidirectional associative memories," *Brain Research*, vol. 1434, pp. 189–199, 2012.

[20] R. M. Gomes, A. P. Braga, and H. E. Borges, "Information storage and retrieval analysis of hierarchically coupled associative memories," *Information Sciences*, vol. 195, pp. 175–189, 2012.

[21] R. Kalaba, Z. Lichtenstein, T. Simchony, and L. Tesfatsion, "Linear and nonlinear associative memories for parameter estimation," *Information Sciences*, vol. 61, no. 1-2, pp. 45–66, 1992.

[22] H. Ushida, T. Yamaguchi, K. Goto, and T. Takagi, "Fuzzy-neuro control using associative memories, and its applications," *Control Engineering Practice*, vol. 2, no. 1, pp. 129–145, 1994.

[23] M. Aldape-Perez, C. Yanez-Marquez, O. Camacho-Nieto, A. J. Arguuelles-Cruz, "An associative memory approach to medical decision support systems," *Computer Methods and Programs in Biomedicine*, vol. 106, no. 3, pp. 287–307, 2012.

[24] C. A. Kumar, M. S. Ishwarya, and C. K. Loo, "Formal concept analysis approach to cognitive functionalities of bidirectional associative memory," *Biologically Inspired Cognitive Architectures*, vol. 12, pp. 20–33, 2015.

[25] W. Pedrycz, "The principle of justifiable granularity and an optimization of information granularity allocation as fundamentals of granular computing," *Journal of Information Processing Systems*, vol. 7, no. 3, pp. 397–412, 2011.

Witold Pedrycz

He is Professor and Canada Research Chair (CRC) in Computational Intelligence in the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Canada. He received M.Sc., Ph.D., and D.Sci. all from the Silesian University of Technology, Gliwice, Poland, in 1977, 1980, and 1984, respectively. He is also with the Systems Research Institute of the Polish Academy of Sciences, Warsaw, Poland. In 2009 Dr. Pedrycz was elected a foreign member of the Polish Academy of Sciences. In 2012 he was elected a Fellow of the Royal Society of Canada. Witold Pedrycz has been a member of numerous program committees of IEEE conferences in the area of fuzzy sets and neurocomputing. In 2007 he received a prestigious Norbert Wiener award from the IEEE Systems, Man, and Cybernetics Society. He is a recipient of the IEEE Canada Computer Engineering Medal, a Cajastur Prize for Soft Computing from the European Centre for Soft Computing, a Killam Prize, and a Fuzzy Pioneer Award from the IEEE Computational Intelligence Society. Current ranking information is available at www.guide2research.com/scientists/CA.