International Journal of Approximate Reasoning 54 (2013) 1149–1167

Contents lists available at SciVerse ScienceDirect

International Journal of Approximate Reasoning

journal homepage: www.elsevier.com/locate/ijar

A granular neural network: Performance analysis and application to re-granulation

Scott Dick *, Andrew Tappenden, Curtis Badke, Olufemi Olarewaju

Department of Electrical & Computer Engineering, University of Alberta, Edmonton, AB, Canada

ARTICLE INFO

Article history:
Received 22 July 2011
Received in revised form 24 October 2012
Accepted 21 January 2013
Available online 28 February 2013

Keywords:
Granular computing
Neuro-fuzzy systems
Neural networks
Fuzzy logic
Linguistic variables
Linguistic hedges

ABSTRACT

The multi-granularity problem is one of the key open problems in Granular Computing. Multiple descriptions of the same phenomena may use very different information granulations, complicating any comparison or synthesis of those descriptions. One method for solving this problem is to transform all observations to a common granulation; however, this granulation must be adequate to capture all important facets of the phenomena. Determining this "natural" granulation could be done by inductively learning and comparing multiple granular representations of the phenomenon, but this requires a dedicated learning architecture. We present the Granular Neural Network, a novel adaptive neural network architecture that employs granular values and operations at the level of individual neurons. The Granular Neural Network is based on the multiplayer perceptron architecture and the backpropagation learning algorithm with momentum. It uses the operations of linguistic arithmetic to manipulate granular connection weights, which are represented by linguistic terms. We test the performance of the Granular Neural Network on three well-known benchmark datasets, and then explore its use in determining the "natural" granularity of a dataset.

1. Introduction

Granular Computing (GrC) is a new approach to computation, which recognizes that in many cases, precision is expensive and unneccesary in modeling the real world [1]. Specifically, GrC attempts to mimic the human ability to reason with abstract quantities and groups rather than numeric precision. It is a framework for reasoning with these abstractions, in order to create computer systems employing a human-centric view of the world [2]. GrC arises from the fusion of several different reasoning formalisms, including fuzzy sets, rough sets [3–6], fuzzy-rough sets [7], interval analysis, shadowed sets [2], and their various hybridizations and generalizations. Type-2 fuzzy sets in particular [8,9], have attracted considerable interest. Applications of granular computing include software reliability modeling [10], Internet search [11], data mining [12], spatial pattern recognition [13], human resource management [14], water management [15], information retrieval [16], and recommender systems [17].

In GrC, the atomic units of data are groups of objects (granules) from some collection. These are aggregated into groups, and then all reasoning and computation is performed on these groups. The principle research issues in GrC revolve around how we create these groups (i.e., *forming* granules), and how we then perform computations using them. In forming granules, we will collect some subset of objects together, based on some measure of "relatedness" (e.g., similarity or indistinguishability); clustering techniques (e.g., [2,18–20]) are one approach for this purpose. However, information granules are not merely clusters; a granule must also be semantically *meaningful* in the context of the application domain. Plainly, this is a matter of interpretation; in some cases, granules are meaningful when they provide predictive power; in others, they

* Corresponding author. Address: Department of Electrical & Computer Engineering, 2nd Flr ECERF Bldg., University of Alberta, Edmonton, AB, Canada T6G 2V4. Tel.: +1 (780) 492 6216; fax: +1 (780) 492 1811.
E-mail address: dick@ece.ualberta.ca (S. Dick).

0888-613X/$ - see front matter
http://dx.doi.org/10.1016/j.ijar.2013.01.012

must be descriptive to a human audience. The key point, however, is that human judgment (in determining what *constitutes* "meaningfulness," as well as in the design of predictive experiments, the assessment of description, etc.) is an important element of forming information granules. Granules might be fuzzy sets, rough sets, intervals, etc.; the specific formalism employed should be congruent to the semantic meaning of the granules [1,21–23].

Once the granules are formed, all computations are performed on them rather than the individual objects. Suggestions on the design of granular operations and algorithms have been put forward in e.g., [24,25], (data fusion via clustering) [26–29], (computing with words) [30,31], (OWA operators), [32] (weighted ordinal means). The idea of "precisiated natural language" and the associated idea of generalized constraints have been advocated in recent years [33–35], with the concept of Z-numbers being a recent extension [36]. Granular operations must again be congruent to the semantics of the information granules. E.g., it is not enough to compare two fuzzy sets for an arbitrary ordering; the semantics of that comparison must be congruent to the semantics of the granules those fuzzy sets represent [1]. Furthermore, human thought appears to employ *many* granulations at once, moving between them seamlessly. Any particular collection of objects is aggregated into an "appropriate" granulation, based on our current perceptions; if those perceptions change, we will revisit that granulation as well. This phenomenon of *multi-granularity* is one of the key sub-problems in designing granular operations, as different information sources with different granulations often have to be combined in real-world applications of GrC (e.g., decision support). This then leads directly to another major sub-problem: how do we represent the results of our computations in terms of the original information granules (the *linguistic approximation*problem)? Finally, the operations may accept granules arising from different formalisms (the *multi-formalism* problem) [22,23,37].

In previous work [28], we have argued that, while humans may use many granulations simultaneously, there is a "natural" granularity to any given problem, and if the information granules are made coarser than this the essential nature of the problem will be lost. Experiments with a synthetic dataset supported this contention. We then demonstrated how *linguistic arithmetic* (which is itself an approach to linguistic approximation) can be used to coarsen or refine information granules, which can help resolve the multi-granularity problem. However, there is currently no way to determine this "natural" granularity *a priori*. One possible solution is to inductively learn granular representations of a problem, determine when these granulations are "adequate," and then retrieve this adequately-granulated knowledge through an explanatory mechanism. We believe that such an approach could be applied to any framework for multi-granular computations (e.g., Herrera's 2-tuple model [37]). However, our ongoing investigation is based on the linguistic arithmetic framework. We select this approach as qualitative behaviors differentiating adequate vs. inadequate granulations were demonstrated for this framework in [28]. The key first step in this project is thus to build an inductive learning system that is congruent to this framework. We can then begin to explore its potential for determining "natural" granularity.

We propose the Granular Neural Network (GNN) architecture, a neural network in which the connection weights (and thus the stored knowledge in the network) are linguistic terms, updated during training via linguistic arithmetic. The GNN (expanding on the initial architecture described in [26,38]) is intended to help explore the "natural" granularity of arbitrary learning problems, by studying how the predictive quality of the network changes as the granularity of the weights is altered. As an obvious prerequisite to using the GNN to study "natural" granularity, we first evaluate this architecture on well-known benchmark problems, comparing its performance against its numerical counterpart, the multi-layer perceptron (MLP) architecture. Despite the much coarser precision of the granular connection weights in the GNN, the overall performance of the GNN was comparable to the MLP architecture in our tenfold cross-validation experiments. As a cross-check on our results, we have also compared these results to the C4.5 decision-tree learner; again, the GNN is comparable to this well-known architecture in our experiments. We then conduct an exploratory investigation of the effect of coarser or refined information granules on the performance of the GNN. Based on the findings in [28], we expect that there will be a threshold effect, wherein the GNN's performance (i.e., out-of-sample prediction accuracy) would markedly decay when the information granularity became too coarse (note that the threshold would be dataset-dependent). We find that this does occur, indicating that the GNN (once coupled with an appropriate explanatory mechanism) is potentially a viable platform for empirically exploring the concept of "natural" granularity.

The remainder of this paper is organized as follows. In Section 2 we review related work on granular computing, granular neural networks, and neuro-fuzzy systems. In Section 3 we review the essential theoretical background on linguistic arithmetic and the original GNN architecture from [26]. In Section 4 we present our improved GNN architecture, including the new backpropagation with momentum learning algorithm. We present our performance analysis experiments in Section 5, and our exploration of different information granularities in Section 6. We close with a summary and discussion of future work in Section 7.

2. Related work

The multi-granularity problem arises because people (even subject-matter experts) often describe the same phenomenon in different ways, and at different levels of granularity (coarser or more refined). We must then establish some sort of common meaning between these different representations of the phenomenon; this has also been called communication between *granular worlds* [2]. Herrera et al. [39] dealt with multigranularity by creating a new, more refined term set and mapping all linguistic values to it; a similar approach (building a new term set and mapping all linguistic values to it) was used in [27,29]. Jiang et al. [40] study a decision-making situation where they transform all linguistic values to fuzzy numbers,