Fuzzy Ant Based Clustering

Steven Schockaert, Martine De Cock, Chris Cornelis, and Etienne E. Kerre

Fuzziness and Uncertainty Modelling Research Unit
Department of Applied Mathematics and Computer Science
Ghent University
Krijgslaan 281 (S9), B–9000 Gent, Belgium
{steven.schockaert,martine.decock,chris.cornelis,etienne.kerre}@UGent.be
http://fuzzy.UGent.be

Abstract. Various clustering methods based on the behaviour of real ants have been proposed. In this paper, we develop a new algorithm in which the behaviour of the artificial ants is governed by fuzzy IF–THEN rules. Our algorithm is conceptually simple, robust and easy to use due to observed dataset independence of the parameter values involved.

1 Introduction

While the behaviour of individual ants is very primitive, the resulting behaviour on the colony-level can be quite complex. A particularly interesting example is the clustering of dead nestmates, as observed with several ant species under laboratory conditions [3]. Without negotiating about where to gather the corpses, ants manage to cluster all corpses into 1 or 2 piles. The conceptual simplicity of this phenomenon, together with the lack of centralized control and a priori information, are the main motivations for designing a clustering algorithm inspired by this behaviour. Real ants are, because of their very limited brain capacity, often assumed to reason only by means of rules of thumb [5]. Inspired by this observation, we propose a clustering method in which the desired behaviour of artificial ants (and more precisely, their stimuli for picking up and dropping items) is expressed flexibly by fuzzy IF–THEN rules.

The paper is organized as follows: in Section 2, we review existing work in the same direction, in particular the algorithm of Monmarché which served as our main source of inspiration. Section 3 familiarizes the reader with important notions about fuzzy set theory and fuzzy IF–THEN rules, while in Section 4 we outline the structure of our clustering algorithm and motivate its key design principles. Some experimental results are presented in Section 5. Finally, Section 6 offers some concluding remarks.

2 Related Work

Deneubourg *et al.* [3] proposed an agent–based model to explain the clustering behaviour of real ants. In this model, artificial ants (or agents) are moving randomly on a square grid of cells on which some items are scattered. Each cell

M. Dorigo et al. (Eds.): ANTS 2004, LNCS 3172, pp. 342–349, 2004.

can only contain a single item and each ant can move the items on the grid by picking up and dropping these items with a certain probability which depends on an estimation of the density of items of the same type in the neighbourhood.

Lumer and Faieta [8] extended the model of Deneubourg *et al.*, using a dissimilarity–based evaluation of the local density, in order to make it suitable for data clustering. Unfortunately, the resulting number of clusters is often too high and convergence is slow. Therefore, a number of modifications were proposed, by Lumer and Faieta themselves as well as by others (e.g. [4, 12]).

Monmarché [10] proposed an algorithm in which several items are allowed to be on the same cell. Each cell with a non–zero number of items corresponds to a cluster. Each (artificial) ant a is endowed with a certain capacity $c(a)$. Instead of carrying one item at a time, an ant a can carry a heap of $c(a)$ items. Probabilities for picking up at most $c(a)$ items from a heap and for dropping the load onto a heap are based on characteristics of the heap, such as the average dissimilarity between items of the heap. Monmarché proposes to apply this algorithm twice. The first time, the capacity of all ants is 1, which results in a high number of tight clusters. Subsequently the algorithm is repeated with the clusters of the first pass as atomic objects and ants with infinite capacity, to obtain a smaller number of large clusters. After each pass k–means clustering is applied for handling small classification errors.

In a similar way, in [6] an ant–based clustering algorithm is combined with the fuzzy c–means algorithm. Although some work has been done on combining fuzzy rules with ant-based algorithms for optimization problems [7], to our knowledge until now fuzzy IF–THEN rules have not yet been used to control the behaviour of artificial ants in a clustering algorithm.

3 Fuzzy IF–THEN Rules

A major asset of humans is their flexibility in dealing with imprecise, granular information; i.e. their ability to abstract from superfluous details and to concentrate instead on more abstract *concepts* (represented by words from natural language). One way to allow a machine to mimic such behaviour, is to construct an explicit interface between the abstract symbolic level (i.e. linguistic terms like "high", "old", ...) and an underlying, numerical representation that allows for efficient processing; this strategy lies at the heart of fuzzy set theory [13], which since its introduction in the sixties has rapidly acquired an immense popularity as a formalism for the representation of vague, linguistic information, and which in this paper we exploit as a convenient vehicle for constructing commonsense rules that guide the behaviour of artificial ants in our clustering algorithm.

Let us recall some basic definitions. A fuzzy set A in a universe U is a mapping from U to the unit interval $[0, 1]$. For any u in U, the number $A(u)$ is called the membership degree of u to A; it expresses to what extent the element u exhibits the property A. A fuzzy set R in $U \times V$ is also called a fuzzy relation from U to V. Fuzzy relations embody the principle that elements may be related to each other to a certain extent only. When $U = V$, R is also called a binary fuzzy