Industrial Marketing Management 42 (2013) 532–543

Contents lists available at SciVerse ScienceDirect

Industrial Marketing Management

Unsupervised KDD to creatively support managers' decision making with fuzzy association rules: A distribution channel application

Albert Orriols-Puig [a], Francisco J. Martínez-López [b,c,*], Jorge Casillas [d], Nick Lee [e]

[a] Grup de Recerca en Sistemes Intelligents, Enginyeria i Arquitectura La Salle, Universitat Ramon Llull, 08022 Barcelona, Spain
[b] Dep. of Business Administration, Business Faculty, University of Granada, Campus Cartuja s/n, 18.071 Granada, Spain
[c] Marketing Group, Dep. of Economic and Business Studies, Open University of Catalonia, Barcelona, Spain
[d] Dep. of Computer Science and Artificial Intelligence, Computer and Telecommunication Engineering School, 18.071 Granada, Spain
[e] Marketing Dep., Aston Business School, Aston Triangle, Birmingham, UK, B4 7ET

ARTICLE INFO

Article history:
Received 24 October 2011
Accepted 16 January 2012
Available online 26 March 2013

Keywords:
Intelligent systems
KDD
Unsupervised learning
Management support
Genetic fuzzy systems

ABSTRACT

To be competitive in contemporary turbulent environments, firms must be capable of processing huge amounts of information, and effectively convert it into actionable knowledge. This is particularly the case in the marketing context, where problems are also usually highly complex, unstructured and ill-defined. In recent years, the development of marketing management support systems has paralleled this evolution in informational problems faced by managers, leading to a growth in the study (and use) of artificial intelligence and soft computing methodologies. Here, we present and implement a novel intelligent system that incorporates fuzzy logic and genetic algorithms to operate in an unsupervised manner. This approach allows the discovery of interesting association rules, which can be linguistically interpreted, in large scale databases (KDD or Knowledge Discovery in Databases.) We then demonstrate its application to a distribution channel problem. It is shown how the proposed system is able to return a number of novel and potentially-interesting associations among variables. Thus, it is argued that our method has significant potential to improve the analysis of marketing and business databases in practice, especially in non-programmed decisional scenarios, as well as to assist scholarly researchers in their exploratory analysis.
© 2013 Elsevier Inc. All rights reserved.

1. Introduction

A core competence for competitive success in current business environments is a firm's capability to process huge amounts of information (Cooper, Watson, Wixom, & Goodhue, 2000), and to generate and then disseminate relevant knowledge to help members of the organization make decisions (Sabherwal & Sabherwal, 2005; Sher & Lee, 2004; Tanriverdi, 2005). To accomplish these aims, a suitable Business Intelligence (BI) system, adequately integrated into a Knowledge Management (KM) process, is of major importance (Alavi & Leidner, 2001; Cody, Kreulen, Krishna, & Spangler, 2002; Herschel & Jones, 2005). Throughout the last few decades, organizations have demanded the development of analytical methods better able to provide added-value information to support strategic and operational decisions (Avison, Eardley, & Powel, 1998; Earl, 2001). In this regard, tools and methodologies that can be used by firms to achieve superior knowledge of their competitive environment are among the most important (Melville, Kraemer, & Gurbaxani, 2004). A widely accepted term to integrate the diversity of tools and computerized systems used by firms for this purpose is *Marketing Management Support Systems* (MMSS, see Van Bruggen & Wierenga, 2001; Wierenga & Van Bruggen, 1993, 1997, 2000). If one takes into account the central position that marketing has in the firm's plan to design and manage their information resources (see Li, McLeod, & Rogers, 2001; Wierenga, van Bruggen, & Althuizen, 2008), advancing MMSS should be of strategic relevance for IS researchers, analysts and managers.

The development of MMSS has been parallel to an evolution in the type of decision problems most pertinent to managers. In particular, while some contemporary marketing problems are well-structured, many others, particularly those within the strategic sphere, are more complex, ill-defined, or unstructured (Wierenga, 2010). Adequately dealing with these latter problems is one of the greatest challenges for today's IS community, and for the development of BI-based solutions (see Negash & Gray, 2008). Such problems frequently require systems able to both work with, and provide information outputs expressed in, qualitative terms, in line with the kind of judgmental process that decision makers have to tackle in these cases. In the last 25 years, this fact has motivated an increasing interest in Artificial Intelligence (AI)-based systems for marketing management support within the general context

* Corresponding author at: Dep. of Business Administration, Business Faculty, University of Granada, Campus Cartuja s/n, 18.071 Granada, Spain. Tel.: +34 958 249595; fax: +34 958 246222.
E-mail addresses: aorriols@gmail.com (A. Orriols-Puig), fjmlopez@ugr.es (F.J. Martínez-López), casillas@decsai.ugr.es (J. Casillas), n.j.lee@aston.ac.uk (N. Lee).

0019-8501/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.indmarman.2013.03.005

of knowledge-driven MMSS. In particular, the recent explosion of interest in Knowledge Discovery in Databases (KDD) methodologies within AI has presented true potential to take the MMSS evolution to a next stage. KDD technologies are a key pillar on which emergent BI systems are based (Petrini & Pozzebon, 2009), though they are both relatively underdeveloped in managerial and research fields at present. For example, Jourdan, Rainer, and Marshall (2008) observe a scarcity in BI research in IS journals during the last decade. Hence, developing KDD applications for information processing and decision-making problems in marketing is a substantial research opportunity. However, the quality of such applications depends on how well described the marketing problem to be solved is, as well as how well developed and applied the AI-based methods are. Consequently, the coordination between KDD specialists and marketing managers is essential to optimize the knowledge generated by KDD applications (see Wang & Wang, 2008).

In the present study, we describe and demonstrate a promising knowledge-driven MMSS category called Marketing *i*ntelligent Systems (see Casillas & Martínez-López, 2010), which bases on KDD and Soft Computing (SC) methodologies for marketing decision-making support. In particular, we present and apply an ad-hoc developed KDD method that we call *Fuzzy-CSar*. The approach combines fuzzy logic (Zadeh, 1994) with a genetic algorithm (Goldberg, 1989; Holland, 1975), resulting in a genetic fuzzy system, within the general parameters of SC methodology (see Zadeh, 1994). The proposed system is based on the use of an evolutionary unsupervised learning method to recover interesting and reliable information patterns (i.e., fuzzy association rules) from databases. In other words, our method is designed to work without *a priori* information on any relationships among the variables contained in the database. The fact that the search process is not driven by a structure of reference (e.g., a model) provides interesting benefits when applied to face new, non-habitual decisional scenarios. Further, the use of fuzzy rules allows users (i.e., managers and analysts) of this method to work with linguistic semantics to define the variables of the database. This particular aspect of the method implies a clear benefit for the kind of information we are interested in (i.e., qualitative), as it allows a linguistic interpretation of the relationships found among the variables defining a certain decision problem. Also, the kind of output offered by this method is similar to the internal rules depicted in human reasoning. We demonstrate our method (i.e. the experimentation) on a distribution channel problem analyzed by Gilliland and Bello (2002).

In order to set the context, we begin with a discussion of MMSS in the context of intelligent systems, followed by short discussions on the use of KDD methodology and intelligent systems as an aid to managerial decision making. We then introduce the proposed method and its unique features, such as the ability to work with multi-item measurements, and qualitative terminology. Subsequently, we demonstrate the potential of the proposed system on the aforementioned distribution channel data set.

2. The evolution of MMSS to Marketing *i*ntelligent Systems

MMSSs have become an increasingly critical part of general Management Information Systems. Their essential purpose is to improve marketing decisions by providing relevant information and reducing uncertainty. Such systems are a key element of an organization's ability to successfully negotiate environmental complexity (Li, Kinman, Duan, & Edwards, 2000). The development of MMSSs from the 1960s has been strongly determined by advances in both information technologies and artificial intelligence in the last two decades (see Li, 2000; Wierenga & Van Bruggen, 1993). Here, a synthetic overview of the main types of MMSS is described, with reference to previous significant contributions in this regard (see also Sisodia, 1992a; Talvinen, 1995; Wierenga & Van Bruggen, 1997, 2000).

The first major classification criteria distinguishes between data-driven and knowledge-driven systems (see Van Bruggen & Wierenga, 2000; Wierenga, Van Bruggen, & Staelin, 1999), in order to better understand the purpose of the MMSS. Data-driven systems are quantitatively-oriented and look for optimal solutions for marketing management problems. The most significant types of such systems are: Marketing Models; Marketing Information Systems (e.g. Brien & Stafford, 1968; Cox & Good, 1967; Kotler, 1966); and Marketing Decision Support Systems (e.g., Keen & Scott-Morton, 1978; Little, 1970, 1975, 1979; Montgomery & Urban, 1970; Sprague & Carlson, 1982).

Such systems require structured quantitative data to work effectively, and use defined rules to optimise solutions. Thus, this kind of system is not well suited to supporting the semi- or deficiently-structure decision scenarios common to turbulent contemporary business environments. This fact motivated an evolution in MMSSs toward the analytical methods emerging from the Artificial Intelligence field, which are more versatile and able to work with both qualitative information and ill-structured problems (Casillas & Martínez-López, 2009). Some relevant examples of what can be called knowledge-driven systems are: Marketing Expert Systems (Abraham & Lodish, 1987; Alpar, 1991; Burke, Rangaswamy, Wind, & Eliashberg, 1990; Dubelaar, Finlay, & Taylor, 1991; McCann & Gallagher, 1990; McDonald & Wilson, 1990; Moutinho, Curry, & Davies, 1993; Sisodia, 1992b); Case-based Reasoning (e.g., Burke, 1991; Kolodner, 1993; McIntyre, Achabal, & Miller, 1993); or Marketing Creativity Support Systems (e.g., Elam & Mead, 1990; MacCrimmon & Wagner, 1994).

In recent years, increasingly sophisticated and complex AI-based information solutions have been developed, indicating a new type of MMSS, termed Marketing *i*ntelligent Systems or Mk*i*S (see Martínez-López & Casillas, 2009). Though these are AI-based systems applied to aid decision-taking in marketing management, and thus belong to the broad category of knowledge-driven systems, the new methodologies they apply for knowledge discovery set them apart from more general knowledge-based systems. Mk*i*S are systems which primarily rely on Soft Computing (SC) methodologies and, to a lesser extent, on other hybridized machine-learning methods, to extract knowledge from marketing-relevant databases. SC is an emerging field of Computer Science and Artificial Intelligence research, which implies cooperative (rather than autonomous) activity for computing paradigms such as fuzzy logic, neural networks, evolutionary computation and probabilistic reasoning (Zeleznikow & Nolan, 2001). The main tenet of SC is exploiting the tolerance of imprecision and uncertainty to achieve robust and low-cost solutions (Zadeh, 1994), with its final aim to provide methods which emulate the human mind to solve complex, real-world problems (Bonissone, 1997). Given the great potential of SC for application in the support of the ill-structured problems typical to marketing and management (Casillas & Martínez-López, 2010), developments in Mk*i*S should be a key issue for IS researchers.

3. KDD as a base for Business Intelligence and Mk*i*S development

Because of their strong impact on a firm's competitive advantage, companies should be involved with effectively managing their information resources (Piccoli & Ives, 2005). The firm's decision process, and thus competitive position, is strengthened by improvements in the process of moving data to information to knowledge (see Bharadwaj, 2000; Gold, Malhotra, & Segars, 2001). Consequently, the key question of how best to increase performance in this area has motivated firms to continuously explore the potential of IT advances, and knowledge technologies in particular (Sambamurthy, Bharadwaj, & Grover, 2003). One of the most promising lines is the application of KDD-based methods to aid in solving management and marketing problems (Bose & Mahapatra, 2001; Buckman, 2004; Cui, Wong, & Lui, 2006; Feng & Chen, 2007; Herschel & Jones, 2005). KDD is a research field focused on the identification of potentially-valid patterns in data (Fayyad, Piatesky-Shapiro, Smyth, & Uthurusamy, 1996), and it can be considered as one of the key BI technologies (Weidong, Weihui, &

Kunglong, 2010). Moreover, as any successful KDD process requires the participation of human knowledge, it should be a natural connector between BI and KM in organizations (see Wang & Wang, 2008). In applying KDD, it is necessary to develop a multi-stage process, starting from the preparation of data in tractable format in the machine learning stage, and finishing with the interpretation of outputs and the assimilation of new knowledge by the firm.

The structure depicted in Fig. 1 is appropriate when adapting the KDD process for business applications (Cabena, Hadjinian, Stadler, Verhees, & Zanasi, 1998; Han & Kamber, 2001): (1) identification and problem delimitation; (2) data preparation (pre-processing); (3) data mining (machine learning); (4) analysis, evaluation and interpretation of results; and (5) presentation, assimilation and use of knowledge. The data mining stage is considered the core of the KDD process. Data mining is characterized by the application of machine-learning methods to automatically or semi-automatically extract comprehensible and useful patterns or models from data (Witten & Frank, 1999, 2000). The Fuzzy-CSar method described subsequently bases the machine learning stage on a particular SC methodology, termed a genetic fuzzy system.

Nevertheless, even considering the significant role of data mining in the KDD process, it would be a mistake to overlook the importance of the application of the prior and subsequent stages (Fayyad & Simoudis, 1995). In essence, the whole process should be oriented toward the elicitation of useful knowledge to be disseminated through, and applied by, the organization. In other words, the output (i.e., new knowledge) obtained after the application of this BI supporting solution has to be adequately integrated into the KM process of companies (see Herschel & Jones, 2005).

4. Unsupervised KDD to creatively aid managerial decision making

A typical managerial practice is to base the analysis and extraction of relevant information from commercial databases on models (i.e., data/variables relational structures). Thus, the analytical methods applied to obtain information *condition* their process of search on an a priori structure that it is supposedly subjacent to the data. Nevertheless, in some cases, the use of models to drive a knowledge-search process might imply some limitations. In particular, unlike the typical scholarly research situation, where in accordance with the scientific method, analytical processes are highly conditioned by theoretical frameworks, managers usually subordinate such issues to the generation of high-quality information in order to face specific situational decision problems. Taking this final aim as driving tenet, decision-makers might thus be prone to disregard the use of a priori models when analyzing data. As noted above, managers have to face certain problems which are not properly structured or delimitated, and/or where there is no *a priori* information to orient analysis. Here, the analytical tools usually applied to routine, well-programmed problems are not adequate. Such problems demand more creative solutions (see Wierenga & Van Bruggen, 2000), so managers must be open-minded when assessing the application of new methods. In this regard, there are many artificial intelligence/knowledge-based methods with potential, when properly designed and adapted, to achieve good performance in non-structured, ill-defined problems. However, such methods are underused at present to support decisions in marketing management (Wierenga, 2010). Even so, as noted above, the evolution of MMSS has shown a clear tendency toward strengthening the role of knowledge-based methods. In fact, these, and KDD-based methods in

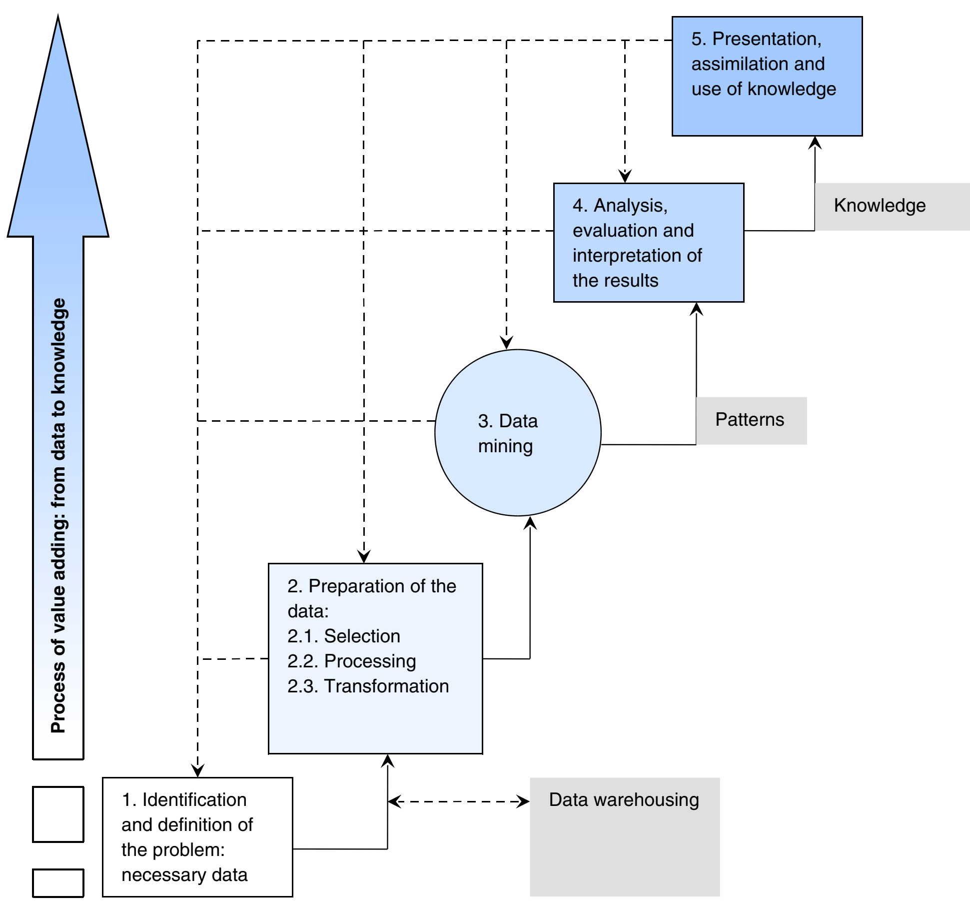

Fig. 1. A business-applied view of the KDD process.

particular, offer the best solutions for the kind of marketing problems discussed herein.

When applying a KDD process, two main approaches can be distinguished for what is termed the 'machine learning' stage (Zhang, 2004): a top-down approach, where the expert takes as a base an *a priori* relational structure for the set of variables defining the problem or database to be analyzed (i.e., supervised learning); or a bottom-up approach (i.e., non-supervised learning), where there is no existing information on what structures may connect the data. The latter approach is based on methods that automatically explore the data in order to find out eventual subjacent information patterns (i.e., connections between variables). Such methods have a philosophy of application different to most IS studies (the context in which machine learning work is most popular), where researchers (or managers) usually start with a theoretical framework to set and test hypotheses (or instead, a relational structure among variables in a business database) by means of a supervised machine learning technique (Michael & Paul, 1999).

At first sight, the outputs provided by a bottom-up KDD process might seem more difficult to work with, as one may expect managers must need to discard a lot of information patterns, many of them describing irrelevant or spurious relationships between variables. In this regard, it is important to have in mind the type of marketing problems of interest, i.e., undefined or ill-structured problems, with little solid theoretical foundations. These problem characteristics make the application of model-driven analytical methods more difficult, if not impossible. However, information patterns extracted from a database after an exploratory, unsupervised machine learning method can be helpful to understand such problems (Padmanabhan & Tuzhilin, 2002), assisting the manager to take better decisions. In this regard, a KDD method designed with a bottom-up approach might also stimulate the creativity of users – see for example the ORAC models described by Wierenga and Van Bruggen (1997) – as it motivates a divergent reasoning. In other words, it is open to many likely options, even considering apparently non-valid or counter-intuitive solutions, when compared with a more orthodox, convergent reasoning process.

5. Fuzzy association rules to assist with the analysis of marketing databases

Fuzzy association rules mining (Agrawal, Imielinski, & Swami, 1993) is a type of unsupervised learning technique that aims to extract interesting, useful and novel information from data, thus inducing the aforementioned divergent managerial reasoning. Three of the main characteristics of these techniques, which make them very appealing for the kind of problems marketing managers face, are that they (1) are able to automatically discover interesting associations among variables or constructs, without requiring any a priori information, (2) use linguistic terms in the rules, allowing a kind of reasoning which is analogous to humans, and (3) are based on solid fuzzy logic theory (Zadeh, 1994). In what follows, we introduce the basics of fuzzy association rules and discuss how they can help analyze marketing problems.

In order to formally define what association rules are, let us start with the following assumptions. Let $T = \{t_1, t_2, \ldots, t_n\}$ be a set of transactions, where each transaction consists of a set of items $I = \{i_1, i_2, \ldots, i_k\}$. For simplicity, let us assume that an item can be present or not in a transaction, but that quantities are not considered in this first definition. Let an *itemset* X be a collection of items $I = \{i_1, i_2, \ldots, i_m\}$. Then, an association rule R is an implication of the form X → Y, where both X and Y are itemsets with no intersecting items. An example of association rule in a shopping environment could be:

if the customer purchases product 1 **then** he/she also purchases product 2

The goal of association rule mining is then to discover the most interesting association rules. While there exist numerous ways to measure the interest of association rules, the two basic indicators of the quality of the rules are *support* (supp) and *confidence* (conf). The support of the rule indicates the frequency of occurring patterns, that is:

$$\mathrm{supp}(\mathrm{A}{\rightarrow}\mathrm{B}) = \frac{1}{N}|\{t{\in}T|(A{\cup}B)\ t\}|,$$

where N is the number of transactions in the database and A and B are the itemsets in the antecedent and the consequent of the association rule.

Confidence evaluates the strength of the implication denoted in the association rule as

$$conf(A- > B) = \left|\frac{\{t{\in}T|(A{\cup}B)\ t\}|}{|\{t{\in}T|A\ t\}|.}\right.$$

It is interesting to analyze the meaning of both support and confidence from a statistical point of view. Support is an exact measure of the frequency with which the items in the rule appear together in the transactions of the database; therefore, it is the probability that the items in A and B appear together in a transaction. The confidence can be interpreted as an estimate for the conditional probability of having the itemset A in a transaction *t*, while also having the itemset B in *t*.

The integration of fuzzy logic with association rules came with the need to process transactions whose items were defined by quantitative values (Webb, 2001). This is usually necessary in marketing and many other contexts. Among the different proposals to solve such problems, fuzzy logic provided one of the most appealing solutions since it introduced a descriptive language, enabling the use of linguistic terms to describe each problem variable. This new scenario permitted the extraction of rules such as

if the customer purchases a *low* quantity of product 1 **then** he/she purchases a *high* quantity of product 2

In this example, *low* and *high* are two linguistic variables based on fuzzy set theory and fuzzy logic (Zadeh, 1994). Note that, in a marketing decision context, the use of fuzzy logic in association rules has three key benefits: (1) it allows the use of data sets containing continuous variables, which is most often the case for marketing data; (2) it lets the system represent knowledge with linguistic terms, which can be easily understood by human experts; and (3) the user is able to define the linguistic terms for each variable, permitting a high flexibility of configuration. Regarding this last point, Fig. 2 shows an example of a possible fuzzy partition for the variable *quantity*. Notice in the example that a quantity of 30 belongs to the fuzzy set labelled as *low* with a degree of 0.25, to the fuzzy set labelled as *medium* with a degree of 0.75, and to the fuzzy set labelled as *high* with a degree of 0.

With the redefinition of the rules, the concepts of support and confidence must be slightly changed. Now, support is computed as

$$\mathrm{supp}(\mathrm{A}{\rightarrow}\mathrm{B}) = \frac{1}{N}\sum_{e=1}^{N}\mu_A(x^e)\cdot\mu_B(y^e),$$

where $\mu_A(x^e)$ and $\mu_B(y^e)$ are the membership degree of the input values with the antecedent and the consequent variables respectively, and N is the number of transactions in the database. Note that the support of a rule increases with the number of transactions in the database that have a high membership degree with the variables of both the antecedent and the consequent of the rule. Therefore, the measure denotes the frequency of occurrence of the relationship captured by the rule.

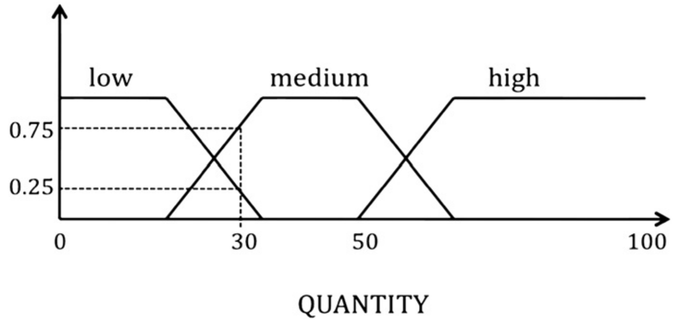

Fig. 2. Illustrative example of a linguistic variable quantity, which is composed of three linguistic terms – i.e., low, medium and high – and its corresponding fuzzy sets. Note that a quantity of 30 has a membership degree of 0.25 to *low* and of 0.75 to *medium*.

On the other hand, confidence is computed as

$$conf(A \rightarrow B) = \frac{\sum_{e=1}^{N} \left(\mu_A(x^e) \cdot \max\{1-\mu_A(x^e), \mu_B(y^e)\}\right)}{\sum_{e=1}^{N} \mu_A(x^e)}.$$

In this case, note that confidence accounts for the frequency in which examples that have a high matching degree with the antecedent of the rule, also have a high matching degree with the consequent of this rule. This measure therefore denotes the strength of the association between the variables in the antecedent and consequent of the rule.

Fuzzy rule mining algorithms aim at extracting fuzzy rules – whose variables are described by linguistic terms – with high support and confidence; that is, rules that denote strong and frequent relationships between the variables in the antecedent and consequent of the rule. In the next section we present Fuzzy-CSar, a novel algorithm that uses these concepts to extract useful knowledge from marketing data.

6. A proposed method for unsupervised learning

The intelligent decision support system presented here, which we have termed Fuzzy-CSar, is an unsupervised learning technique that combines genetic algorithms (Goldberg, 1989; Holland, 1975) with a fuzzy logic representation to extract fuzzy association rules from a set of examples. The aim of the algorithm is to provide a useful tool that not only may help marketing experts to check hypotheses previously established, but can also be used as a source of inspiration to formulate new hypotheses or to approach certain decision problems from new perspectives. This is of particular interest in problems that may not be well structured, and in situations where there is a lack of a priori information to orient analysis.

Below, we briefly introduce the architecture of our system. Subsequently, we introduce the knowledge representation of the system, present the novel concept of multi-item fuzzification – which enables the system to deal with the particularities of much marketing data – and describe the procedure used by the system to automatically discover new and interesting association rules. Our proposal is an evolution of a recently presented learning classifier system designed to extract quantitative association rules from unlabelled data streams dynamically, quickly and efficiently; for further details on the latter system, the reader is referred to Orriols-Puig and Casillas (2010). Our contribution in this paper is to allow the system to extract knowledge more understandable to humans, by using fuzzy logic. This issue is crucial to enable the use of our method by applied researchers and managerial decision makers. To do so, the knowledge representation, the corresponding coding scheme, the inference process, and the genetic operators are newly designed. Further, the proposed Fuzzy-CSar method is also endowed with a mechanism able to deal with multi-item data sets, such as those habitually used to measure unobserved variables (i.e. constructs) in business and other social science research.

6.1. Knowledge representation

The proposed system evolves a *population* of association rules, which are usually referred to as *classifiers*. At the end of the learning process, the population is expected to contain rules that capture the most interesting associations between problem variables. The maximum size of the population is specified by the user. With this, the user sets an upper bound on the number of interesting associations that can be discovered from data.

Each classifier consists of a fuzzy association rule and a set of parameters. The fuzzy association rule is represented as: **if** x_i *is* A_i *and … and* x_j *is* A_j **then** x_c *is* A_c, where the antecedent contains ℓ_a input variables $x_i, \ldots, x_j$ ($0 < \ell_a < \ell$, where ℓ is the number of variables of the problem) and the consequence consists of a single variable x_c which is not present in the antecedent. Each variable has a *linguistic term* or *label*, A_i, assigned from all the linguistic terms that the given variable can take. Note that this structure allows a number of variables to be antecedents, but requires a single variable in the consequent. With this strategy, we are searching for sets of variables with certain values that may cause another one to occur. These types of rules, thus, could be interpreted as a causal relationship between certain values of the variables in the antecedent and certain values of the consequent variable. Support and confidence are used to reflect the interestingness of the rule component of the classifier, computed as indicated in Section 5, with the aim to uncover rules with high support and confidence.

6.2. Multi-item fuzzification

Fuzzy-CSar uses the concept of multi-item fuzzification first proposed by Martínez-López and Casillas (2009), specifically designed to deal with the multi-item latent construct measurement commonly used in marketing and social sciences research. This method assumes that each individual item provides partial information about the corresponding construct (i.e., an unobserved variable or first-order variable). Therefore, once can compute the matching degree as the aggregation (T-conorm) of the information given by each item. Thence, the matching degree of a variable i with the vector of items $x_i = (x_1{}^i, x_2{}^i, \ldots, x^i_{pi})$ is

$$\mu_{A_i}(x_i) = \max_{h_i=1}^{p_i} \mu_{A_i}\left(x^i_{p_i}\right),$$

where $\mu_{Ai}(x^i{}_{pi})$ is the matching degree of the variable i represented by the linguistic term Ai with the input $x^i{}_{pi}$. In our experiments, detailed below, we used the maximum as the union operator, implemented as a sum bounded to 1. However, the method is also able to work with variables whose measurement has been done by just one item or input.

6.3. Learning process

Our algorithm incrementally learns from a stream of examples; at each learning iteration, it receives an input example $(e_1, e_2, \ldots, e_{\ell})$ and takes action to incrementally update the classifier's parameters and discover new promising rules. First, the system creates the *match set* [M] with all the classifiers in the population that match the input example with a degree larger than 0. If [M] contains less that θ_{mna} classifiers, the *covering operator* is triggered to create as many new matching classifiers as required to have θ_{mna} classifiers in [M]. Then, classifiers in [M] are organized in *association set candidates*. Each association set candidate is given a probability to be selected that is proportional to the average confidence of the classifiers that belong to this association set. The selected *association set* [A] goes through a *subsumption* process which aims to diminish the number of rules that express similar associations among variables. Then, the parameters of all the classifiers in [M] are updated. At the end of the iteration, a genetic algorithm is applied to [A] if the average time since its last application is greater than θ_{GA}. This process is repeatedly applied, therefore, updating the parameters of existing classifiers and creating new promising rules online.

To completely understand how the system works, five elements need further explanation: (1) the covering operator, (2) the procedure to create association set candidates, (3) the association set subsumption mechanism, (4) the parameter update procedure, and (5) the rule discovery by means of a genetic algorithm. In the following subsections, each of these elements is explicated in more detail.

6.3.1. Covering operator

Given the sampled input example *e*, the covering operator creates a new classifier that matches *e* with maximum degree. That is, for each variable, the operator randomly decides (with probability $1 - P_{\#}$) whether the variable has to be in the antecedent of the rule, with the constraints (1) that, at least, a variable has to be selected and (2) that, at most, $\ell - 1$ variables can be included in the antecedent. Then, one of the remaining variables is selected to be in the rule. Each of these variables is initialized with the linguistic label that maximizes the degree of match with the corresponding input value. In addition, we introduce generalization by permitting the addition of any other linguistic term with probability $P_{\#}$, with the restrictions (1) that each variable in the antecedent contains *maxLabIn* linguistic terms at maximum and (2) that each variable in the consequent contains *maxLabOut* linguistic terms at maximum.

6.3.2. Creation of association set candidates

The aim of creating association set candidates, or niches, is to group rules that express similar associations in order to establish competition among them, letting the best ones take over their niche. For this purpose, our system uses the following approach, which relies on the idea that rules that have the same variable with the same or similar linguistic terms in the consequent must belong to the same niche, since probably they would denote similar associations among variables. First, it sorts the rules of [M] ascendantly, depending on the variable of the consequent. Given two rules r_1 and r_2 that have the same variable in the consequent, we consider that r_1 is smaller than r_2 if $\ell_1 < \ell_2$ or ($\ell_1 = \ell_2$ and $u_1 > u_2$), where ℓ_1, u_1, ℓ_2, and u_2 are the position of first and the last linguistic terms of the output variable of each rule respectively.

Once [M] has been sorted, the association set candidates are built as follows. At the beginning, an association set candidate [A] is created and the first classifier in [M] is added to this association set candidate. Then, the following classifier *k* is added if it has the same variable in the consequent, and ℓ_k is smaller than the minimum u_i among all the classifiers in the current [A]. This process is repeated until finding the first classifier that does not satisfy this condition. In this case, a new association set candidate is created, and the same process is applied to add new classifiers to this association set.

6.3.3. Association set subsumption

We designed a subsumption mechanism with the aim of reducing the number of different rules that express the same knowledge. The process works as follows. Each rule in [A] is checked for subsumption with each other rule in [A]. A rule r_i is a candidate subsumer of r_j if it satisfies the following four conditions: (1) r_i has higher confidence and it is experienced enough (that is, $conf^i > conf_0$ and $exp^i > \theta_{exp}$, where $conf_0$ and θ_{exp} are user-set parameters); (2) all the variables in the antecedent of r_i are also present in the antecedent of r_j (r_j can have more variables in the antecedent than r_i); (3) both rules have the same variable in the consequent; (4) r_i is more general than r_j. A rule r_i is more general than r_j if all the input and the output variables of r_i are also defined in r_j, and r_i has, at least, the same linguistic terms as r_j for each one of its variables.

6.3.4. Parameter update

At the end of each learning iteration, the parameters of all the classifiers that belong to the match set are updated. First, the experience of the classifier is incremented. Second, the support of each rule is updated as

$$\text{supp}_{t+1} = \frac{\text{supp}_t \cdot (\ell time - 1) + \tilde{\mu}_A\left(x^{(e)}\right) \cdot \tilde{\mu}_B\left(y^{(e)}\right)}{\ell time},$$

where $\ell time$ is the life time of the classifier, that is, the number of iterations that the classifier has been in the population. Then, the confidence is computed as $\text{conf}_{t+1} = \text{sum\_imp}_{t+1}/\text{sum\_mat}_{t+1}$, where

$$sum\_imp_{t+1} = sum\_imp_t + \tilde{\mu}_A\left(x^{(e)}\right) \cdot \max\left\{1 - \tilde{\mu}_A\left(x^{(e)}\right), \tilde{\mu}_B\left(x^{(e)}\right), \quad and \right.$$
$$sum\_mat_{t+1} = sum\_mat_t + \tilde{\mu}_A\left(x^{(e)}\right).$$

Next, the fitness of each rule in [M] is computed as $F = conf^{v}$, where *v* permits controlling the pressure toward highly fit classifiers. Finally, the association set size estimate of all rules that belong to [A] is updated. Each rule maintains the average size of all the association sets in which it has participated.

6.3.5. Discovery component

The genetic algorithm is triggered on [A] when the average time from its last application upon the classifiers in [A] exceeds the threshold θ_{GA}. It selects two parents p_1 and p_2 from [A], where each classifier has a probability of being selected proportional to its fitness. The two parents are crossed with probability P_{χ}, generating two offspring ch_1 and ch_2. Fuzzy-CSar uses a uniform crossover operator that contemplates the restriction that any offspring has to have, at least, a variable in the rule's antecedent. If crossover is not applied, the children are an exact copy of the parents. The resulting offspring may go through three different types of mutation: (1) mutation of antecedent variables (with probability $P_{I/R}$), which randomly chooses whether a new antecedent variable has to be added to or one of the antecedent variables has to be removed from the rule; (2) mutation of the linguistic terms of the variable (with probability P_{μ}), which selects one of the existing variables of the rule and mutates its value; and (3) mutation of the consequent variable (with probability P_C), which selects one of the variables of the antecedent and exchanges it with the variable of the consequent.

Thereafter, the new offspring are introduced into the population. If the population is full, excess classifiers are deleted from [P] with a probability directly proportional to their association set size estimate, and inversely proportional to their fitness.

7. Application of the proposed system to a distribution channel problem: description and methodology

In order to highlight the benefits that the application of our system to marketing data can provide, a KDD bottom-up approach is used to analyze a marketing channel problem, previously studied by using a traditional scholarly theory-driven approach. Below, the problem considered here is firstly presented as a case study and then the methodology employed is detailed. The next section analyzes the results.

7.1. Study of the database

7.1.1. Brief description

The case study we have used to demonstrate the proposed M*ki*S is described in Gilliland and Bello (2002, hereafter referred to as GB) and it focuses on the antecedents to, and consequences of, attitudinal commitment between manufacturers and independent intermediaries in a distribution channel. To address this problem, GB first defined a structural model illustrated in Fig. 3. It regards attitudinal commitment among channel members as consisting of two components: *calculative commitment* and *loyalty commitment*. *Calculative commitment* is "the state of attachment to a partner cognitively experienced as a realization of the benefits sacrificed and losses incurred if the relationship were to end" (GB p. 28). *Loyalty commitment* is "the state of attachment to a partner experienced as a feeling of allegiance and faithfulness" (GB p. 28).

GB's model assumes that a *manufacturer's loyalty commitment* to the distributor is positively associated with the *use of the social enforcement mechanism* (H1a: +) and negatively associated with the *use of contractual enforcement mechanism* (H1b: −). Conversely, a *manufacturer's calculative commitment* is positively associated with the *use of the contractual enforcement mechanism* (H2a: +) and negatively associated with the *use of the social enforcement mechanism* (H2b: −). In addition, the model assumes that there are three constructs which are positively associated with *manufacturer's calculative commitment:* the *manufacturer's relative dependence* (H3: +), the *manufacturer's pledge of exclusivity* (H4: +), and the *manufacturer's pledges of investments* (H5: +). On the other hand, there are two constructs positively associated with *manufacturer's loyalty commitment* to the distributor: *manufacturer's perceptions of distributor's pledges of investments* (H6: +) and *manufacturer's trust in distributor* (H7: +).

To collect the data, 529 survey packets were mailed to managers carefully selected from firms that produced manufactured products that were sold in industrial applications, from which 314 usable questionnaires were returned. The constructs were partially described by a different number of items. *Manufacturer's relative dependence* was partially measured by three items graded by thirteen-point rating scales. *Manufacturer's pledge of exclusivity* was measured by an item that could take three categorical values. All the other constructs where measured by either three or four items graded by seven-point rating scales. Structural equation modeling was employed to test each of the hypotheses above. All the hypotheses where supported except for hypothesis H1b.

7.1.2. Limitations of theory-driven approaches and research opportunity

While it is clear that GB's approach is an exemplary example of the traditional scientific approach to business research, it is worth pointing out three potential limitations of this theory-driven approach. First, the approach is conditioned by an *a priori* relational structure of the database. As these hypotheses specify only a limited number of relationships between variables, some key relationships may have not been discovered. Second, all the hypotheses only consider relations between pairs of variables, while there could be interesting associations between more than two variables. Third, the statistical analysis tests the correlation between all the values that two variables can take, but is unable to detect whether there are positive or negative associations between some ranges of values of the variables. Further, the result of the statistical analysis is an estimate and a t-value, and the given hypothesis is supported or rejected depending on whether the t-value is greater than or smaller than a given threshold. Note that this type of result provides poor explanation of *why* a hypothesis is rejected or supported; on the other hand, fuzzy association rules carefully explain for which range of values the association is found and give an idea its frequency (support) and strength (confidence).

In order to extend this analysis, in the following we use Fuzzy-CSar to identify associations among the variables of the problem with the aim of detecting any relationship not captured by the initial hypotheses. The approach is not proposed as an alternative to the classical theory-driven process of knowledge generation, but as a complement. However, as previously treated (see Section 4), for the case of managerial applications, such approach could be also used as a primary alternative in decisional scenarios deficiently structured, where there is no a priori information.

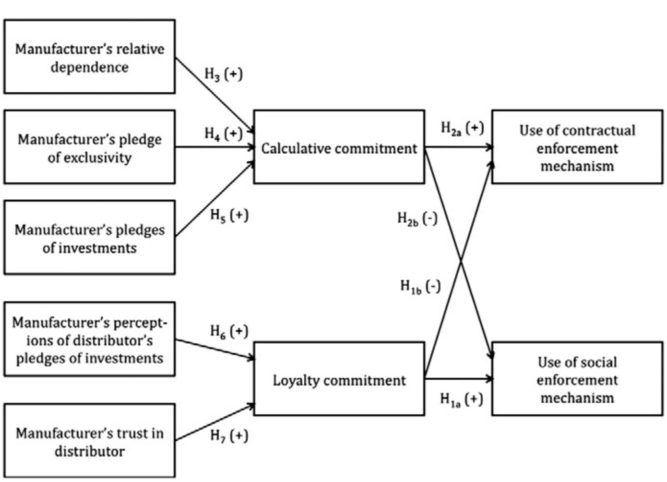

Fig. 3. Conceptual model of attitudinal commitment, its antecedents and consequences proposed by Gilliland and Bello (2002).

7.2. Experimental methodology

The aim of our experiments was: (1) to study the robustness of the proposed system to deal with marketing data; (2) to examine the ability of the system to support the conclusions obtained by applying a theory-driven approach; (3) to show the capacity of the system to discover new relationships between variables and for specific ranges of variables, which may help better understand the true associations between the problem constructs; and (4) to demonstrate the ease of use of the system. For this purpose, we ran Fuzzy-CSar on the GB's data with no further processing and without informing the system of any a priori structure. The system was configured with a population size of 6400 rules and the following parameters: $P_{\#} = 0.5$, $P\chi = 0.8$, $\{P_{I/B}, P_{\mu}, P_C\} = 0.1$, $\theta_{GA} = 50$, $\theta_{exp} = 1\,000$, $conf_0 = 0.95$, $\nu = 1$, $\delta = 0.1$.

The fuzzy partition for each item was defined as follows. For all items graded by thirteen-point rating scales and seven-points rating scales, we used three linguistic terms defined by 2 trapezoidal-shaped membership functions for the first and third linguistic term and a triangular-shaped membership function for the second linguistic term, as shown in Fig. 4a and b respectively. We used this partition since each linguistic term has the same area. Finally, for the item that describes *manufacturer's pledge of exclusivity*, which could take three different values, we employed three linguistic terms defined by triangular-shaped membership functions, each one centered on one of the three possible values, as shown in Fig. 4c.

8. Analysis of the results

In this section, we report the results of the experiments detailed above and compare the conclusions to those obtained in GB's original work. We first examine whether the current methodology supports the hypotheses tested by GB. Second, we show some further useful, interesting and novel information extracted with a KDD bottom-up approach. While the latter information illustrates the kind of outputs that our system can provide to support ill-structured decisional problems in marketing, its utility can also be extended to many other areas.

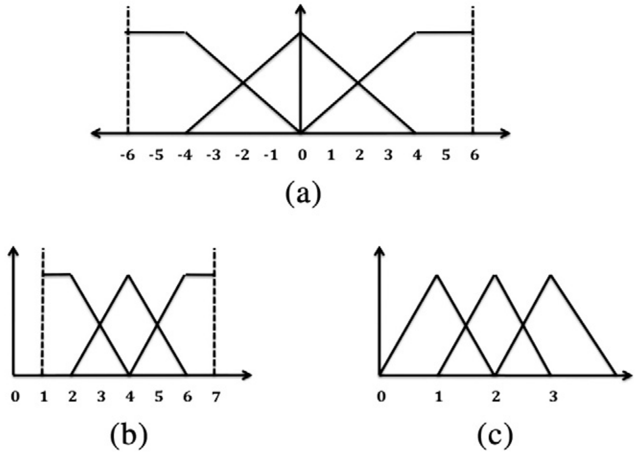

Fig. 4. Fuzzy partitions defined for items described by (a) thirteen-point rating scales, (b) seven-point rating scales, and (c) three categorical values.

8.1. Does Fuzzy-CSar supports the original model hypotheses?

To check whether the association rules extracted by our system support GB, we first searched in the final populations for rules with high support and confidence that had only one variable in the antecedent and another one in the consequent, therefore defining a simple association as done by GB. Although we considered all the rules, we were particularly interested in rules in which the variables in the antecedent and in the consequent had the same linguistic term or had completely opposite linguistic terms. That is to say, considering the three linguistic terms employed – i.e., *small* (S), *medium* (M), and *large* (L) – we searched for rules such as "*if* var$_i$ *is* S *then* var$_j$ *is* S" or rules such as "*if* var$_i$ *is* S *then* var$_j$ *is* L." The first case corresponds to a positive relationship between the variables, whilst the second case corresponds to a negative relationship between variables.

The rules evolved by Fuzzy-CSar led us to the following conclusions from each hypothesis (in the following rules, *s* and *c* stand for support and confidence):

There were two fuzzy association rules that supported GB's hypothesis 1a:

a) *High loyalty commitment → high use of social enforcement mechanism* with s = 0.603 and c = 0.943.
b) *Medium loyalty commitment → medium use of social enforcement mechanism* with s = 0.351 and c = 0.556.

These two rules indicate a positive association between *loyalty commitment* and *use of social enforcement*. In addition, note that the use of linguistic terms enables us to get a better understanding of the association between both variables. That is, the most frequent and strong association can be found for *high* values of both variables; rule a) denotes that there is a high number of examples that support the association and that almost any example with *high loyalty commitment* resulted in a *high use of social enforcement mechanism* (the confidence is almost 1). On the other hand, *medium* values of both variables resulted in a frequent and strong association, but not as frequent and strong as for *high* values. In addition, the system could not evolve any rule for *low* values of both variables. All this indicates that the two variables are particularly associated when they take *medium* and, especially, *high* values.

There were three rules that provided interesting information about GB's hypothesis 1b:

a) *High loyalty commitment → low use of contractual enforcement mechanism* with s = 0.412 and c = 0.673.
b) *Medium loyalty commitment → low use of contractual enforcement mechanism* with s = 0.428 and c = 0.663.
c) *Medium loyalty commitment → medium use of contractual enforcement mechanism* with s = 0.413 and c = 0.633.

The two first rules denote a strong association between *high* and *medium* values of *loyalty commitment* and *low* values of *use of contractual enforcement mechanism.* The third rule indicates that there is a strong relationship between *medium* values of both variables. These three variables highlight that there is not a negative association between both variables as concluded by the structural equation modeling approach of GB. Nevertheless, note that our system provides us with much more information; it can detect particular associations between ranges of values of both variables. Rule a) indicates that there indeed exists a negative association between both variables, but only for *high* values of *loyalty commitment* and *low* values of *use of contractual enforcement mechanism*. However, there is not a clear negative association between the other ranges of values of both variables. Rules b) and c) indicate that *medium* values of *loyalty commitment* yielded *low* and *medium* values of *use of contractual enforcement mechanism*, which does not support hypothesis 1b.

There were two rules that supported GB's Hypothesis 2a:

a) *Medium calculative commitment → medium use of contractual enforcement mechanism* with s = 0.317 and c = 0.632.
b) *Low calculative commitment → low use of contractual enforcement mechanism* with s = 0.389 and c = 0.675.

The two rules denote that a manufacturer's calculative commitment to the distributor is positively associated with the use of the contractual enforcement mechanism, although this association can be found only for *medium* and *low* values of both variables.

There were two rules that strongly supported Hypothesis 2b in GB:

a) *Low calculative commitment → high use of social enforcement mechanism* with s = 0.506 and c = 0.842.
b) *Medium calculative commitment → high use of social enforcement mechanism* with s = 0.458 and c = 0.865.

Both rules indicate that a manufacturer's calculative commitment to the distributor is negatively associated with the use of the social enforcement mechanism. Note that the most frequent association appeared for *low* values of *calculative commitment* and *high* values of *use of social enforcement mechanism.*

There were three rules that strongly supported GB's hypothesis 3:

a) *Medium manufacturer's relative dependence → medium calculative commitment* with s = 0.411 and c = 0.577.
b) *Low manufacturer's relative dependence → low calculative commitment* with s = 0.248 and c = 0.846.

These two rules therefore indicate a positive association between *manufacturer's relative dependence* and *calculative commitment*. In addition, the rules also show that the most frequent association can be found for *medium* values of both variables; but at the same time, *medium* values led to not-so-strong associations (a confidence value of 0.577). On the other hand, the strongest association can be found for *low* values of both variables, although this situation is not as frequent as having *medium* values.

There was a single association rule that supported hypothesis 4 in GB:

a) *Low manufacturer's pledge of exclusivity → low calculative commitment* with s = 0.470 and c = 0.642.

More specifically, the rule denotes that a *low* value of *manufacturer's pledge of exclusivity to the distributor* is associated with a *low* value of *manufacturer's calculative commitment to the distributor*. Fuzzy-CSar did not extract supported association rules for *medium* and *high* values of these variables, which indicates that the correlation between both variables can only be found for *low* values of these two variables.

There were two rules that supported GB's hypothesis 5:

a) *Medium* manufacturer's pledges of investments → *medium* calculative commitment with s = 0.342 and c = 0.574.
b) *Low* manufacturer's pledges of investments → *low calculative commitment* with s = 0.278 and c = 0.729.

That is, these two rules indicate that *low* and *medium* values of *manufacturer's pledges of investments to the distributor* are positively associated with, respectively, *low* and *medium* values manufacturer's *calculative commitment to the distributor*.

There were three rules that supported GB's hypothesis 6:

a) *High manufacturer's perceptions of distributor's pledges of investments → high loyalty commitment* with s = 0.372 and c = 0.792.
b) *Medium manufacturer's perceptions of distributor's pledges of investments → medium loyalty commitment* with s = 0.445 and c = 0.725.

The two rules indicate that *medium* and *high* values of manufacturer's perceptions of the distributor's pledges of investments are positively associated with, respectively, *medium* and *high* values of manufacturer's loyalty commitment to the distributor.

Finally, there were three rules that supported hypothesis 6 in GB:

a) *High manufacturer's trust in distributor → high loyalty commitment* with s = 0.438 and c = 0.791.
b) *Medium manufacturer's trust in distributor → medium loyalty commitment* with s = 0.381 and c = 0.756.

That is, *high* and *medium* values of manufacturer's trust in the distributor are positively associated with, respectively, *high* and *medium* values of manufacturer's loyalty commitment to the distributor.

The overall analysis conducted in this subsection has not only confirmed the results obtained by GB, but also highlighted the value added by the application of the fuzzy-association-rule mining approach. Specifically, the use of linguistic terms has enabled the use of a kind of reasoning which is very similar to the human one. In addition, it has increased the understanding of where the strong and frequent associations are across the range of the continuous variables — which may allow a more detailed appreciation of relationships between variables. Therefore, this case study has highlighted the benefits that the application of Fuzzy-CSar can yield, which can be used as a complement of more traditional theory-driven statistical techniques such as structural equation modelling. Next, we further emphasize the utility of this system by showing the new interesting variable associations which were uncovered by our analysis.

8.2. A bottom-up approach to discover new/unexpected relationships in data sets

One of the key characteristics of our system is that it can automatically discover new interesting associations directly from data. The system can be especially useful to discover relationships overlooked by marketing experts in situations such as an improper structure or delimitation of the problem or a lack of a priori information to orient the original analysis. Thus, we report here the most interesting fuzzy association rules discovered by our system that were not considered in GB's original analysis.

Rules that indicate a direct relationship between two variables that are mediated by a third variable in the original structural model:

a) *Low manufacturer's pledge of exclusivity → high use of social enforcement mechanism* with s = 0.619 and c = 0.846.
b) *High manufacturer's trust in distributor → high use of social enforcement mechanism* with s = 0.537 and c = 0.948.
c) *High manufacturer's pledges of investments → high use of contractual enforcement mechanism* with s = 0.503 and c = 0.911.

Rules that denote associations not considered in GB's original structural model:

a) *Medium loyalty commitment → medium manufacturer's relative dependence* with s = 0.533 and c = 0.783.
b) *Low use of contractual enforcement mechanism → high use of social enforcement mechanism* with s = 0.529 and c = 0.851.
c) *Medium manufacturer's pledges of investments → medium manufacturer's relative dependence* with s = 0.497 and c = 0.789.

While it is beyond the present scope to deal in detail with each one of these potential hypotheses, delving into the conceptual discussion presented in GB suggests that there is some theoretical merit in further examining such relationships. More specifically, a manufacturer's pledge of exclusivity mandates exclusive rights to a single distributor, raising the financial and temporal stakes of dissolving the relationship. Without such a pledge, manufacturers are likely to feel that they need to place more emphasis on less formal modes of governance (i.e. social enforcement), quite apart from the commitment-based mediating mechanism argued by GB. High trust in a distributor implies that the manufacturer perceives the distributor as benevolent, and likely to deliver on its promises (i.e. credibility). It also indicates shared expectations and values between the two organizations. It seems very likely that high levels of trust will lead a manufacturer to be far more comfortable in employing informal modes of governance like social enforcement directly. Conversely, if a manufacturer makes significant pledges of investment in a distributor, they are placing valuable financial and time resources into that relationship (e.g. providing sales material, training, etc.). In such cases, manufacturers will naturally be keen to formally protect their potential return on such investments, by utilizing formal, contractual based modes of enforcement. While GB argue that all three of the above relationships are mediated fully by the two forms of commitment, they do not formally test this mediation, and there is conceptual merit in examining these direct links.

Regarding the second set of new association rules, again there looks to be some merit in further examining them. The association between medium loyalty commitment and medium relative dependence suggests an interesting mechanism not considered by the theory used by GB. In this situation, perhaps the emotional bond of loyalty commitment leads in some situations to the manufacturer's feeling that they have the most to lose if the relationship is dissolved. High loyalty may lead a manufacturer to ‘turn down’ other partnership options (or potential partners to avoid approaching), leading to a lack of viable options for the manufacturer, which is a key component of relative dependence. It is also interesting to consider the potential link between the two different enforcement mechanisms. In fact, it makes good sense to argue that a lower level of formal contractual enforcement will naturally leave a gap in governance which the manufacturer will be keen to make up with less formal social enforcement mechanisms. Finally, it stands to reason that if manufacturers invest in a distribution partner (pledges of investment) to some extent, they will feel they have more to lose with relationship dissolution (i.e. relative dependence).

Furthermore, Fuzzy-CSar is able to uncover non-hypothesized relationships between more than two variables, which is not possible without *a priori* modeling in regression and SEM. For example, rules that denote very strong associations among more than two variables (which could be considered as similar to interactive relationships in a regression context) uncovered from GB's sample were:

a) *Low manufacturer's pledge of exclusivity* and *high loyalty commitment → high use of social enforcement mechanism* with s = 0.438 and c = 0.943
b) *High manufacturer's trust in distributor* and *high loyalty commitment → high use of social enforcement mechanism* with s = 0.427 and c = 0.981

c) *High loyalty commitment* and *low use of contractual enforcement mechanism → high use of social enforcement mechanism* with s = 0.387 and c = 0.949.

Again, in a theoretical sense, such rules appear to have enough merit to warrant further investigation and confirmatory testing in a scholarly context. For example, while GB's original H1a argued for a link between high loyalty commitment and high social enforcement, in situations where there is a low level of manufacturer exclusivity, the link may be especially strong. This is because in such contexts, there is little constraint on the manufacturer with regard to using competitive distributors, and it is likely that the only thing bonding the manufacturer to a distributor in any kind of channel relationship at all is some kind of informal social enforcement mechanism, which is naturally stronger in a situation where the manufacturer feels some kind of emotional bond (loyalty) to the distributor. The link between loyalty and social enforcement is also likely to be especially strong in situations of high trust, since trust and loyalty may operate as a kind of self-reinforcing system or 'virtuous circle' — each helping to build the other. Finally, it makes some sense to argue that the link between low levels of contractual enforcement and high levels of social enforcement will be stronger with higher levels of loyalty commitment, since the greater interaction in loyal partnerships will lead to a higher level of cooperative behaviors, each reinforcing the willingness of the partners to rely on such mechanisms to govern the relationship in the absence of contracts.

All these rules indicate the existence of new and potentially interesting associations between variables, that not previously considered by knowledge extraction methods based on a priori structures (i.e. models). In fact, existing research in the area has not examined in detail the associations posited above — perhaps because much extant channels theory structures conceptual development in such a way that researchers may avoid hypothesizing such relationships. Thus, the marketing expert can then read the information provided by these kinds of rules and use them to explain or support a decisional problem from new perspectives.

9. Conclusions

In this paper, we have highlighted the importance of marketing intelligent systems – and in particular, of computerized techniques based on AI – in order to extract novel, useful and interesting knowledge from large volumes of data that may be helpful to support strategic and operational decisions. More specifically, we have proposed the use of Fuzzy-CSar, a KDD solution that does not need to have any a priori information about the problem, to complement the theory-driven analysis of marketing problems and data bases. The value of the new methodology can be leveraged by any marketing problem in general, but especially for the complex, ill-defined or unstructured problems that are usually found in strategic spheres.

In order to test the value provided by the proposed methodology, we have addressed the marketing scenario originally treated in GB. We have applied our system to the original data without any a priori information on potential subjacent structures of relationships (e.g. possible models) in variables, and have analyzed the system outcome – i.e., the fuzzy rules returned by the system – at the end of the run. The results suggest the following four key issues.

- The application of Fuzzy-CSar supported all the hypotheses also supported by the multivariate statistical technique applied in the parameterization stage of the original study (SEM, in this case); the system therefore yielded similar conclusions as the theory-driven approach in the first instance.
- In addition to confirming the original hypothesis, the proposed system provided an explanation of why each hypothesis was either supported or rejected.
- As our system does not depend on any a priori model, it could discover new interesting associations between two or more variables that integrate certain decisional scenario.
- The use of fuzzy logic – and specifically, of linguistic terms – resulted in a type of association rules that could be easily read by subject experts, and also by management practitioners, especially when compared to the probabilistic values returned by the statistical tests usually applied in traditional scholarly analyses.

Thus, it can be seen that the unsupervised learning approach looks to have substantial utility for practitioners looking for robust methods with which to interrogate the large data bases commonly found in the marketing function of an organization.

Another interesting point to note when discussing the utility of our approach in this regard concerns the characteristics of the structural equation modelling method used in GB. This is a statistical-based modeling method commonly used in scholarly research, so applying our method also presents benefits within an academic context. If one is to move from a confirmatory approach to SEM to a more exploratory one (e.g. generating hypotheses from the data), one could use various indications – such as modification indices – provided by the SEM package to explore the potential for additional hypotheses for testing. However, the advantage of Fuzzy-CSar here is clear in the rules that were reported above. Specifically, our system can uncover rules concerning only part of the range of each variable – e.g. medium loyalty commitment associated with medium relative dependence. Such a rule would be unlikely to return a positive modification index in SEM because SEM in general requires linear relationships, and therefore needs a relationship to be relatively consistent across the entire range of the dependent and independent variables. Thus, relying on SEM to provide information on such additional relationships is potentially dangerous. While this may not be a problem in the testing of a strong *a priori* theory, if one looks to directly test issues such as competing hypotheses (of which mediation is one example), or other more exploratory goals, Fuzzy-CSar looks to have much to offer.

Overall, the results of our case study indicate that the application of the Marketing Intelligent System we propose in this paper promises great benefits, especially with respect to the ease of its application. Likewise, we would like to highlight (particularly to scholars) that the application of Fuzzy-CSar does not necessarily need to be taken as a replacement of the typical theory-driven approach, but as a complement. That is to say, the system can be applied to raw data as a complement to test hypotheses initially specified by a theory-driven approach and to detect new interesting relationships, or to serve as inspiration to the marketing expert to tackle decisional problems from new positions.

Acknowledgments

The authors would like to thank Ministerio de Educación y Ciencia for its support under projects TIN2008-06681-CO6-01 and TIN2008-06681-CO6-05, Generalitat de Catalunya for its support under grant 2005SGR-00302, and Junta de Andalucía for its support under project P07-TIC-3185. Further, we would like to thank David I. Gilliland and Daniel C. Bello for their kind consent to our use of their database.

References

Abraham, M. M., & Lodish, L. M. (1987). PROMOTER: An automated promotion evaluation system. *Marketing Science*, *6*(2), 101–123.

Agrawal, R., Imielinski, T., & Swami, A. (1993). Mining association rules between sets of items in large databases. *Proceedings of the ACM SIGMOD International Conference on Management of Data* (pp. 207–216).

Alavi, M., & Leidner, D. E. (2001). Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS Quarterly*, *25*(1), 107–136.

Alpar, P. (1991). Knowledge-based modelling of marketing managers´ problem solving behaviour. *International Journal of Research in Marketing*, *8*, 5–16.

Avison, D. E., Eardley, W. A., & Powel, P. (1998). Suggestions for capturing corporate vision in strategic information systems. *Omega*, *26*(4), 443–459.

Bharadwaj, A. S. (2000). A resource-based perspective on information technology capability and firm performance: An empirical investigation. *MIS Quarterly*, *24*(1), 169–196.

Bonissone, P. (1997). Soft Computing: The convergence of emerging reasoning technologies. *Soft Computing - A Fusion of Foundations, Methodologies and Applications*, *1*(1), 6–18.

Bose, I., & Mahapatra, R. K. (2001). Business data mining: A machine learning perspective. *Information Management*, *39*(3), 211–225.

Brien, R. H., & Stafford, J. E. (1968). Marketing information systems: A new dimension for marketing research. *Journal of Marketing*, *32*(3), 19–23.

Buckman, R. H. (2004). *Building a Knowledge-Driven Organizations.* New York, NY: McGraw Hill.

Burke, R. R. (1991). Reasoning with empirical marketing knowledge. *International Journal of Research in Marketing*, *8*(1), 75–90.

Burke, R. R., Rangaswamy, A., Wind, J., & Eliashberg, J. (1990). A knowledge-based system for advertising design. *Marketing Science*, *9*, 212–229.

Cabena, P., Hadjinian, P., Stadler, R., Verhees, J., & Zanasi, A. (1998). *Discovering Data Mining: From Concept to Implementation.* : Prentice Hall.

Casillas, J., & Martínez-López, F. J. (2009). Mining uncertain data with multiobjective genetic fuzzy systems to be applied in consumer behaviour modeling. *Expert Systems with Applications*, *36*(2, part 1), 1645–1659.

Casillas, J., & Martínez-López, F. J. (Eds.). (2010). *Marketing intelligent systems using soft computing: Managerial and research applications..* : Springer.

Cody, W. F., Kreulen, J. T., Krishna, V., & Spangler, W. S. (2002). The integration of business intelligence and knowledge management. *IBM Systems Journal*, *41*(4), 697–713.

Cooper, B. L., Watson, H. J., Wixom, B. H., & Goodhue, D. L. (2000). Data warehousing supports corporate strategy at first American corporation. *MIS Quarterly*, *24*(4), 547–567.

Cox, D. F., & Good, R. E. (1967). How to build a marketing information system. *Harvard Business Review*, *45*(3), 145–154.

Cui, G., Wong, M. L., & Lui, H. (2006). Machine learning for direct marketing response models: Bayesian network with evolutionary programming. *Management Science*, *52*(4), 597–612.

Dubelaar, C., Finlay, P. N., & Taylor, D. (1991). Expert systems: The cold fusion of marketing? *Journal of Marketing Management*, *7*(4), 371–382.

Earl, M. (2001). Knowledge management strategies: Toward a taxonomy. *Journal of Management Information Systems*, *18*(1), 215–233.

Elam, J. J., & Mead, M. (1990). Can software influence creativity? *Information Systems Research*, *1*(1), 1–22.

Fayyad, U. M., Piatesky-Shapiro, G., Smyth, S., & Uthurusamy, R. (1996). *Advances in knowledge discovery and data mining.* : M.I.T. Press.

Fayyad, U. M., & Simoudis, E. (1995). Knowledge Discovery and Data Mining: Tutorial. *Proceedings of the 14th international joint conference on artificial intelligence, Montreal, Canada*.

Feng, D., & Chen, E. T. (2007). Firm performance effects in relations to the implementation and use of knowledge management systems. *International Journal of Innovation and Learning*, *4*(2), 172–185.

Gilliland, D. I., & Bello, D. C. (2002). Two sides to attitudinal commitment: The effect of calculative and loyalty commitment on enforcement mechanisms in distribution channels. *Journal of the Academy of Marketing Science*, *30*(1), 24–43.

Gold, A. H., Malhotra, A., & Segars, A. H. (2001). Knowledge management: An organizational capabilities perspective. *Journal of Management Information Systems*, *18*(1), 185–214.

Goldberg, D. E. (1989). *Genetic algorithms in search, optimization & machine learning.* : Addison Wesley.

Han, J., & Kamber, M. (2001). *Data Mining. Concepts and Techniques.* : Morgan Kauffmann Publishers.

Herschel, R. T., & Jones, N. E. (2005). Knowledge management and business intelligence: The importance of integration. *Journal of Knowledge Management*, *9*(4), 45–55.

Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems.* Cambridge, MA: MIT Press.

Jourdan, Z., Rainer, R. K., & Marshall, T. E. (2008). Business Intelligence: An Analysis of the Literature. *Information Systems Management*, *25*, 121–131.

Keen, P. G. W., & Scott-Morton, M. S. (1978). *Decision support systems: An organizational perspective.* Reading, MA: Addison-Wesley.

Kolodner, J. (1993). *Case-based reasoning.* San Mateo, CA: Morgan Kauffman.

Kotler, P. (1966). A design for the firm's marketing nerve center. *Business Horizons*, *9*, 63–74.

Li, S. (2000). The development of a hybrid intelligent system for developing marketing strategy. *Decision Support Systems*, *27*(4), 395–409.

Li, S., Kinman, R., Duan, Y., & Edwards, J. S. (2000). Computer-based support for marketing strategy development. *European Journal of Marketing*, *34*(5/6), 551–575.

Li, E. Y., McLeod, R., Jr., & Rogers, J. C. (2001). Marketing information systems in Fortune 500 companies: A longitudinal analysis of 1980, 1990, and 2000. *Information Management*, *38*(5), 307–322.

Little, J. D. C. (1970). Models and managers: The concept of a decision calculus. *Management Science*, *16*(8), B466–B485.

Little, J. D. C. (1975). BRANDAID: A marketing mix model, Part I: Structure; Part II: Implementation. *Operations Research*, *23*, 628–673.

Little, J. D. C. (1979). Decision support systems for marketing managers. *Journal of Marketing*, *43*, 9–26.

MacCrimmon, K. R., & Wagner, C. (1994). Stimulating ideas through creativity software. *Management Science*, *40*(11), 1514–1532.

Martínez-López, F. J., & Casillas, J. (2009). Marketing Intelligent Systems for consumer behaviour modelling by a descriptive induction approach based on Genetic Fuzzy Systems. *Industrial Marketing Management*, *38*(7), 714–731.

McCann, J. M., & Gallagher, J. P. (1990). *Expert systems for scanner environments*. Boston: Kluwer Academic Publichers.

McDonald, M. H. B., & Wilson, H. N. (1990). State-of-the-art development in expert systems and strategic marketing planning. *British Journal of Management*, *1*(3), 159–170.

McIntyre, S. H., Achabal, D. D., & Miller, C. M. (1993). Applied case-based reasoning to forecasting retail sales. *Journal of Retailing*, *69*, 372–398.

Melville, N., Kraemer, K., & Gurbaxani, V. (2004). Review: Information technology and organizational performance: An integrative model of IT business value. *MIS Quarterly*, *28*(2), 283–322.

Michael, C. H., & Paul, G. (1999). Special section: Data mining. *Journal of Management Information Systems*, *16*(1), 11–16.

Montgomery, D. B., & Urban, G. L. (1970). Marketing decision information systems: An emerging view. *Journal of Marketing Research*, *7*, 226–234.

Moutinho, L., Curry, B., & Davies, F. (1993). The COMSTRAT model: The development of an expert system in strategic planning. *Journal of General Management*, *19*(1), 32–47.

Negash, S., & Gray, P. (2008). Business intelligence. In F. Burstein, & C. W. Holsapple (Eds.), *Handbook on Decision Support Systems*, *2*. (pp. 175–193): Springer.

Orriols-Puig, A., & Casillas, J. (2010). Evolution of interesting association rules online with learning classifier systems. In J. Bacardit, W. Browne, J. Drugowitsch, E. Bernadó-Mansilla, & M. V. Butz (Eds.), *Learning Classifier Systems, Lecture Notes in Computer Science*, *Volume 6471*. (pp. 21–37): Springer.

Padmanabhan, B., & Tuzhilin, A. (2002). Knowledge refinement based on the discovery of unexpected patterns in data mining. *Decision Support Systems*, *33*(3), 309–321.

Petrini, M., & Pozzebon, M. (2009). Managing sustainability with the support of business intelligence: Integrating socio-environmental indicators and organisational context. *The Journal of Strategic Information Systems*, *18*, 178–191.

Piccoli, G., & Ives, B. (2005). Review: IT-Dependent strategic initiatives and sustained competitive advantage: A review and synthesis of the literature. *MIS Quarterly*, *29*(4), 747–776.

Sabherwal, R., & Sabherwal, S. (2005). Knowledge management using information technology: Determinants of short-term impact on firm value. *Decision Sciences*, *36*(4), 531–567.

Sambamurthy, V., Bharadwaj, A., & Grover, V. (2003). Shaping agility through digital options: Reconceptualizing the role of information technology in contemporary firms. *MIS Quarterly*, *27*(2), 237–263.

Sher, P. J., & Lee, V. C. (2004). Information technology as a facilitator for enhancing dynamic capabilities through knowledge management. *Information Management*, *41*(8), 933–945.

Sisodia, R. S. (1992a). Marketing information and decision support systems for services. *Journal of Services Marketing*, *6*(1), 51–64.

Sisodia, R. S. (1992b). Expert marketing with expert systems. *Marketing Management*, *1*(2), 32–47.

Sprague, R. H., & Carlson, E. D. (1982). *Building effective decision support systems.* Englewood Cliffs, NJ: Prentice Hall.

Talvinen, J. M. (1995). Information systems in marketing: Identifying opportunities for new applications. *European Journal of Marketing*, *29*(1), 8–26.

Tanriverdi, H. (2005). Information technology relatedness, knowledge management capability and performance of multibusiness firms. *MIS Quarterly*, *29*(2), 311–334.

Van Bruggen, G. H., & Wierenga, B. (2000). Broadening the perspective on marketing decision models. *International Journal of Research in Marketing*, *17*, 159–168.

Van Bruggen, G. H., & Wierenga, B. (2001). Matching management support systems and managerial problem-solving modes: The key to effective decision support. *European Management Journal*, *19*(3), 228–238.

Wang, H., & Wang, S. (2008). A knowledge management approach to data mining process for business intelligence. *Industrial Management & Data Systems*, *108*(5), 622–634.

Webb, G. I. (2001). Discovering associations with numeric variables. *Proceedings of the seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 383–388).

Weidong, Z., Weihui, D., & Kunglong, Y. (2010). The relationship of Business Intelligence and knowledge management. *Proceedings of the 2nd IEEE International Conference on Information Management and Engineering (ICIME)* (pp. 26–29).

Wierenga, B. (2010). Marketing and artificial intelligence: Great opportunities, reluctant partners. In J. Casillas, & F. J. Martínez-López (Eds.), *Marketing intelligent systems using soft computing: Managerial and research applications* (pp. 1–8). Springer.

Wierenga, B., & Van Bruggen, G. T. (1993). Marketing management support systems: Evolution, adoption and effectiveness. In ESOMAR/EMAC/AFM (Ed.), *Information based decision making in marketing: Reconciling the needs and interests of decision makers, data collectors and data providers* (pp. 1–20).

Wierenga, B., & Van Bruggen, G. T. (1997). The integration of marketing problem-solving modes and marketing management support systems. *Journal of Marketing*, *61*, 21–37.

Wierenga, B., & Van Bruggen, G. T. (2000). *Marketing management support systems: Principles, tools and implementation.* : Kluwer Academic Publishers.

Wierenga, B., van Bruggen, G. H., & Althuizen, N. A. P. (2008). Advances in marketing management support systems. In B. Wierenga (Ed.), *Handbook of Marketing Decision Models. International Series in Operations Research & Management Science*. (pp. 561–592). : Springer.

Wierenga, B., Van Bruggen, G. T., & Staelin, R. (1999). The success of marketing management support systems. *Marketing Science*, *18*(3), 196–207.

Witten, I. H., & Frank, E. (1999). *Tools for data mining: practical machine learning tools and techniques with Java implementations.* : Morgan Kauffman.

Witten, I. H., & Frank, E. (2000). *Data mining: practical machine learning tools and techniques with Java implementations.* San Francisco, CA, USA: Morgan Kaufmann Publishers.

Zadeh, L. A. (1994). Soft computing and fuzzy logic. *IEEE Software*, *11*, 48–56.

Zeleznikow, J., & Nolan, J. R. (2001). Using soft computing to build real world intelligent decision support systems in uncertain domains. *Decision Support Systems*, *31*(2), 263–285.

Zhang, M. (2004). *Introduction to Data Mining and KDD Process.* Technical report, Victoria University of Wellington (Available at: http://www.mcs.vuw.ac.nz/courses/COMP422/2004T2/Lecturenotes/week01.pdf)

Albert Orriols-Puig. He received the M.Sc. and Ph.D. degrees in computer engineering in 2004 and 2008, respectively, from the Ramon Llull University, Spain. His thesis studied how the extended classifier system (XCS), one of the most influential LCS, could deal with domains that contained class imbalances. In 2009, he was appointed as an assistant professor at the Ramon Llull University. In 2010, he took a software engineer position at Google. His research interests include online evolutionary learning, fuzzy modeling, learning from rarities, data complexity, and machine learning in general. He is especially interested in the application of genetic-based machine learning to real-world problems in the field of supervised and unsupervised learning. He serves as a reviewer for several conference and machine learning journals.

Francisco J. Martínez-López is a Professor of Business Administration at the University of Granada and the Open University of Catalonia (Barcelona) in Spain. He has been visiting scholar in the marketing department at the Zicklin School of Business (CUNY, USA), Aston Business School (Aston University, UK), or the Michael Smurfit School of Business (University College Dublin, Ireland), among others. He is the Editor-in-Chief of the International Journal of Business Environment (Inderscience Publishers). Dr. Martínez-López has co-edited several international journals' special issues and research books for leading publishers of business and management research. Likewise, he has published more than 60 papers in international conferences proceedings and journals.

Jorge Casillas is an Associate Professor with the Department of Computer Science and Artificial Intelligence, University of Granada (Spain), where he is a member of the Soft Computing and Intelligent Information Systems research group. He has authored a book, edited four international books, edited three international journal special issues, and organized seven special sessions in international conferences on topics such as interpretability-accuracy trade-off in fuzzy modeling, genetic fuzzy systems, genetic-based machine learning, intelligent robotics, and marketing intelligent systems. He is author of about 30 journal papers, 12 book chapters, and 60 conference papers. He serves on the Editorial Board of the Evolutionary Intelligence journal of Springer since January, 2008. His research interests include fuzzy modeling, intelligent robotics, marketing intelligent systems, knowledge discovery, and metaheuristics.

Nick Lee is Professor of Marketing and Organizational Research, and the Research Convenor, for the Marketing Group. His main research interests all flow from the domains of social/organizational psychology and research methods. He is the Editor-in-Chief of the European Journal of Marketing. He has won multiple awards for his work. In 2009 he was featured in The Times as 'one of the 15 scientists whose work will shape the future'. His work has appeared or is forthcoming in journals such as the Journal of the Academy of Marketing Science, the Journal of Business Research, the Journal of Business Ethics, Industrial Marketing Management, the European Journal of Marketing, the Journal of Strategic Marketing, and the American Journal of Bioethics, among others.