Available online at www.sciencedirect.com

ScienceDirect

Fuzzy Sets and Systems 285 (2016) 6–30

FUZZY sets and systems

www.elsevier.com/locate/fss

On generating linguistic descriptions of time series

Nicolás Marín [a,*], Daniel Sánchez [a,b]

[a] Department of Computer Science and Artificial Intelligence, University of Granada, 18071 – Granada – Spain
[b] European Centre for Soft Computing, Mieres – Spain

Received 12 September 2014; received in revised form 24 March 2015; accepted 27 April 2015

Available online 4 May 2015

Abstract

In this paper we provide a general approach for the concepts and processes related to the generation of linguistic descriptions of time series. As we will see, this approach consists of two main tasks, namely, a knowledge extraction task, which can be seen as a Knowledge Discovery in Databases (KDD) procedure, and a linguistic expression process. The presented approach incorporates as a core element a description model, which is based on three pillars: a knowledge representation formalism, an expression language, and a quality framework. In the paper, we also analyze the main tools and techniques that can be used regarding the mentioned tasks and pillars of the generation of linguistic descriptions of time series. Additionally, we provide a deep review of the main contributions in the area, which come mainly from the fields of Natural Language Generation (NLG) and Fuzzy Sets and Systems. The existing and potential contributions of fuzzy sets and extensions are discussed in detail. Together with the application of KDD techniques, we encourage the cooperation of the Fuzzy Sets and the NLG communities in order to provide a significant step forward in the development of systems for providing linguistic descriptions of time series data.

1. Introduction

ICTs professionals have to face the challenge of providing users with useful knowledge from an ever-growing flood of data, most usually for decision support purposes. Very frequently, data come in the form of time series, i.e., in a broad sense, sequences of data regarding a given phenomenon that are ordered in time. For instance, knowing the electricity consumption pattern through the year is crucial for the consumer in order to choose a certain contract rate and for adapting his/her habits, thus saving both money and the natural environment. An even more routine example is that of weather forecasting data, containing relevant and useful knowledge for a wide range of users with different information needs. Beyond these, as we shall see, there is a large amount of applications in economy, healthcare, and social welfare, among others, where having regular access to information contained in time series data is a must.

Time series data are usually provided to users in the form of tables and graphical representations. In order to obtain knowledge from the series, the user has to afford an analysis process which is often very complex and hard to

* Corresponding author.
E-mail addresses: nicm@decsai.ugr.es (N. Marín), daniel@decsai.ugr.es (D. Sánchez).

http://dx.doi.org/10.1016/j.fss.2015.04.014

complete. This is specially the case for non-expert users which, being the final target of the information in most of the cases (as in the examples mentioned in the previous paragraph), lack the necessary background and skills [1].

At the same time, the development of human–computer interaction systems based on natural language, already important in the last decades, is growing in importance nowadays. The wide use of mobile devices together with the many possibilities offered by human-like communication, are behind the development of systems like Apple's SIRI, Google Glass, etc.

A natural language based solution for obtaining information from time series data has been provided in the last fifteen years by the so-called *Data-to-text systems* [2]. They successfully play the role of extracting useful knowledge automatically from the data on the basis of the necessary expert knowledge, using at the same time a feasible communication channel for most of modern devices. Quoting Reiter et al. [3]:

> Data-to-text systems are motivated by the belief that (brief) linguistic summaries of datasets may in some cases be more effective than more traditional presentations of numeric data, such as tables, statistical analyses, and graphical visualizations (even simple visual/graphical displays require relatively complex cognitive processing (Carpenter and Shah, 1998) [4]). Also linguistic summaries can be delivered in some contexts where visualizations are not possible, such as text messages on a mobile phone, or when the user is visually impaired (Ferres et al., 2006) [5].

Data-to-text systems are originated in the NLG (Natural Language Generation) field, which is the area of Computational Linguistics that is concerned with the mapping from non-linguistic to linguistic expressions [6]. These systems aim at generating texts from non-linguistic (usually numerical) input data, texts which are *linguistic descriptions of data*. Generally speaking, a linguistic description of data expresses knowledge extracted from the data, using natural language. *Generating linguistic descriptions of data* consists of two main tasks: a *knowledge extraction* process which in a broad sense can be considered as KDD (Knowledge Discovery in Databases [7]) defined as the non-trivial acquisition of knowledge which is novel, potentially useful, and easily understandable from data, and a *linguistic expression process* which enhances the understandability and usefulness of the obtained knowledge by appropriately expressing it using natural language.

The generation of linguistic descriptions of time series (GLiDTS) is completely different from more conventional time series analysis techniques such as segmentation, pattern recognition and extraction, forecasting, etc. As we shall see, these and other techniques can eventually be part of the whole process of generation of linguistic descriptions. As an example, consider the abovementioned case of weather forecasting, where the time series data to describe has been obtained using predictive techniques in an initial step.

More specifically, GLiDTS systems are *computational systems able to simulate the answer of a human expert when asked about what he/she can observe or highlight in the time series*. Both the kind of linguistic descriptions and the techniques for generating them can differ significantly, even for the same data domain. For instance, a different answer is expected to the requests *Tell me about the patient's last night*, in which a detailed description of a whole series of monitoring data is expected, and *Let me know about significant changes of the patient's heart rate*, where the objective is to locate segments of the series which match linguistically defined patterns like *increasing*, *highly decreasing*, etc.

Both in the areas of KDD and NLG it is widely recognized the potential contributions of uncertainty representation techniques, particularly those related to the Fuzzy Set Theory and extensions. In KDD and Machine Learning, fuzzy set theory has the potential to produce models that are more comprehensible, less complex, and more robust, being especially useful for representing "vague" patterns and modeling and processing various forms of uncertain and incomplete information [8]. In NLG, fuzzy set theory plays a key role in filling the semantic gap between data and linguistic terms and expressions, and in dealing with the different kinds of uncertainty inherent to the linguistic expression of knowledge about real world data [9–11], in line with the well-known suitability and high potential of fuzzy set theory for representing the semantics of natural language expressions [12].

As a consequence, in the last years there has been an increasing number of proposals coming from the fuzzy set theory community for addressing the problem of generating linguistic descriptions of data, starting from the work of Yager [13] and, more recently, generating linguistic descriptions of time series data.

In this paper we provide a general approach to the concepts and processes related to generating linguistic descriptions of time series. As we will see, this general approach consists of two main tasks: a KDD procedure and a linguistic expression process. In this model, we pay special attention to three pillars of the generation problem: the knowledge representation formalism, the expression language and the quality framework. Along the paper, we analyze the main tools and techniques that can be used in the generation of linguistic descriptions of time series, deeply revising the main contributions existing in the literature, from the pioneering work of Karen Kukich [14] to the more recent proposals from the NLG and fuzzy set theory communities.

2. Generation of linguistic descriptions of time series

The objective of GLiDTS systems is to simulate experts in generating linguistic descriptions of time series. In addition to experts, two other kind of agents are important in relation to these systems, as we shall see. First, the *target user*, which is the person who will receive the description. Second, the *designer*, which is the person who builds the system and has to deal with both experts and target users in order to make the appropriate design choices for the system to be effective.

As mentioned in the introduction, a linguistic description of a time series is a text that expresses knowledge extracted from the data, the latter being a time series (i.e., a sequence of data points, measured typically at successive points in time spaced at uniform time intervals), or a collection of time series. Hence, in order to provide a linguistic description of this kind, two main tasks have to be accomplished by the GLiDTS system: *extraction of the knowledge to be transmitted*, and *generation of a suitable linguistic expression* for the target user.

It is important to remark that these tasks are equally complex and important in GLiDTS systems. This is the point that makes the difference between GLiDTS systems and other NLG systems having a knowledge base as input instead of data to be analyzed and interpreted. It also distinguishes GLiDTS systems from conventional data analysis systems that focus their interest in knowledge extraction, but do not address human communication issues [2,15].

Due to the nature of the first task, it can be considered as a KDD process, the objective of which is to provide a collection of *messages* that computationally represent the knowledge. For that purpose it is necessary to define a specific *knowledge representation formalism* for representing the semantics of messages. The expressive power of the formalism will define the space of possible messages to search for building the final description.

Another key point is to provide suitable measures of the *quality* of both messages and output text in order to guide the generation process. Quality has many different facets, as we shall see, that we call *quality dimensions*. Each one can be assessed in a different way, much in the same way as the relevance of association rules, a well-known knowledge representation formalism in the area of KDD, which is usually measured by their *support* and *confidence*, corresponding to the dimensions of *significance* and *accuracy*, respectively. Dimensions and their measures will be part of what we call the *quality framework* of the GLiDTS system, that we will discuss in Section 5.

One of the key aspects of both formalism and quality framework is the *abstraction level* of representations, since it affects significantly the human perception of the time series [16] that we want to simulate. Lower abstraction levels correspond to fine-grained descriptions of data, where the role of linguistic expression and the target user are less important. There is in general a whole hierarchy of abstraction levels where, as we get higher in the hierarchy, the linguistic expression and aspects related to the target user such as utility, novelty, understandability, etc. play a key role. This aspect is related to, among other factors, the *degree of granularity* in the representation of linguistic terms and expressions relative to features of time series data, such as time intervals, fuzzy partition of the domain of values, etc.

The knowledge discovered in the extraction tasks could be provided to the target user by means of little elaborated output generated from the obtained messages, as usual in KDD systems. However, since we expect the system to simulate an expert user, the output must be built taking into account not only aspects of syntax and semantics, but also pragmatics issues. Hence, generating the linguistic expression is not just a matter of translating messages, but choosing the translation which guarantees an optimal communication between system and target user [9]. More specifically, the final linguistic expression has to be easily understandable, compatible with the knowledge that the user would express from a direct observation of data, and useful, among other aspects.

As a consequence, it is evident the key role played by the target user, not only when defining the representation formalism, but also in what respects the expression language [17]. The target user has also a key role in the definition of the quality framework, not only in what respects the dimensions and measures related to the formalism, but also those related to the relevance of the linguistic expression [18,19] and in validating the final result.

Fig. 1 shows the GLiDTS process and the abovementioned elements that form what we call *description model*: the knowledge representation formalism, the expression language, and the quality framework. Though different architec-

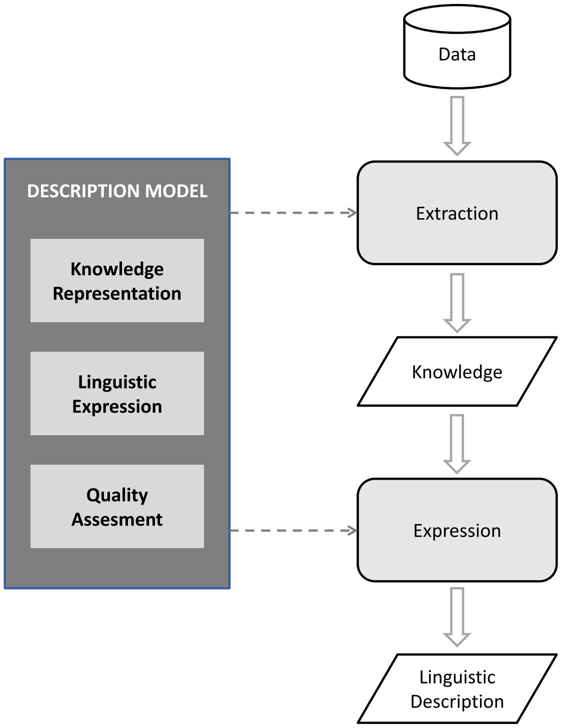

Fig. 1. Components of the GLiDTS process.

tures have been proposed in the literature (see, for example, [14,20,2,21,6]), in all of them the authors highlight the necessity of performing the two tasks (extraction and expression) in Fig. 1.

The rest of the paper analyzes the main components of the presented architecture relying on a thorough review of the literature. Section 3 focuses on the knowledge representation problem in GLiDTS systems, providing a study of formalisms and extraction techniques that can be used for this purpose. Section 4 is devoted to the linguistic expression process, analyzing the different approaches that can be used to face this task, with examples of the kind of outputs that can be found in the existing proposals. Section 5 presents the quality framework as a multidimensional model. Different applications domains where existing GLiDTS systems have been applied are in Section 6. Our conclusions are presented in Section 7. As pointed out in the introduction, fuzzy set theory and related technologies are very useful in both the extraction and expression tasks, and in the specification of the description model. The review of the literature we provide includes different contributions of fuzzy set theory and other Soft Computing techniques in the design and development of GLiDTS systems. We shall summarize these contributions and discuss other potential applications of fuzzy set theory in Section 7.1. Finally, some final reflections are in Section 7.2.

3. Knowledge representation

One of the fundamental aspects of GLiDTS systems, as pointed out in the previous section, is to determine a formalism for representing the knowledge (messages) to be transmitted to the target user. In this paper, we take as general framework the Computing with Words and Perceptions paradigm proposed by L.A. Zadeh, and the closely related Generalized Theory of Uncertainty (GTU) [22]. Though rooted in the realm of fuzzy set theory, it is a suitable approach since knowledge representation formalisms based on classical approaches can be seen as particular cases.

In Zadeh's paradigm, a central role is played by the concept of *protoform*. A protoform (prototypical form) is defined as an abstracted prototype (model). Instances of protoforms represent knowledge about data [23]. Protoforms are crucial for the formalization of human consistent reasoning and deduction capabilities of search, specially in natural language-based knowledge discovery tools [24].[1]

Some specific protoforms have been widely employed in querying and data mining, as well as in GLiDTS systems:

- The protoform *X is A*, where *A* is a linguistic label associated to a fuzzy subset of the domain of a variable *X* (examples of instances are "Temperature is High", "Wind Speed is 33 mph", and "Stock price variation is highly increasing").[2] This is called a *possibilistic constraint* [22].
- The protoform *Q D are A*, where *Q* is a linguistic quantifier, and *D* and *A* are fuzzy subsets of the same universe of reference defined by fuzzy predicates (e.g. "Most days are foggy", "Most times between 12 and 15 h the price is constant").[3] This is a particular case of *fuzzy quantified sentence* [34,22].

However, as we shall see, particular protoforms have been proposed by some authors for specific applications. In Section 3.1, we analyze the definition of the formalism and two related concepts of instance space and semantic space, and we study the main protoforms employed as formalisms for knowledge representation in GLiDTS systems, and in the representation of instance components. Then, in Section 3.2, we focus on the operations that can be performed when carrying out the knowledge extraction task, and we revise both exhaustive and non-exhaustive search techniques within this step, as well as other alternative approaches.

3.1. Formalism

Knowledge representation formalisms are employed in order to represent the semantics of messages through the semantics of protoforms and those of their instance's components, which are the basic elements of any formalism. For example, the semantics of the message represented by the instance "Most times between 12 and 15 h the price is constant" is given by the semantics of:

- the protoform *Q D are A* (the fraction of objects in D that are in A is Q),
- the component *between 12 and 15 h* (represented by a crisp set),
- the component *constant* (which, when measuring the variation of a time series by means of the angle with the horizontal axis of a line adjusted to the series, can be represented by a fuzzy subset of angles around zero),
- the component *Most times* (a fuzzy relative quantifier represented by a non-decreasing fuzzy subset of [0,1])

as well as a collection Ω of time segments (taken from a single series or from different series). Then, D is the (in this case crisp) set of intervals with boundaries 12 and 15 h in Ω, and A is the fuzzy subset of Ω satisfying that the variation of the time series is *constant* in the sense explained before.

We can distinguish two moments when the elements of the formalisms can be defined:

- *A priori*. These are elements of the formalism which are defined without taking into account the actual values of the data. All protoforms fall in this category, as well as those linguistic labels whose representation is provided by the designer of the GLiDTS system, like *Most times* and *constant* in the previous example.
- *A posteriori*. These are elements of the formalism which are obtained by analyzing the time series, searching for linguistic labels as the representation of components, in order to provide the best matching between time series and protoform instances. Many different methods for analysis have been employed in the literature, remarkably

[1] Kacprzyk and Zadrozny point out that protoforms can be seen as a general form of a linguistic data summary, a term coined by Yager for a specific kind of linguistic description based on quantified guided aggregation of data [13], and it is very frequently used as a synonym for linguistic description in the NLG community, see for instance [15].

[2] Notice that in some of these examples *A* is a crisp set, in some cases a singleton, which are particular cases of fuzzy sets. The same can happen with any other component of a protoform instance. Let us stress once more that, since components of protoforms' instances are fuzzy sets, they include crisp expressions as particular cases. In this sense, to the best of our knowledge, they are general enough to cover all the formalisms employed in the literature.

[3] This protoform is deeply related to fuzzy association rules [25,26] and the related fuzzy approximate dependencies [27,28], gradual dependencies [29–32], and exception/anomalies in rules [33], that are examples of other kind of protoforms, coming from the fuzzy data mining field, that have been employed or could be employed in GLiDTS systems.

machine learning (clustering, etc.) and pattern recognition techniques with different degree of complexity, among others. Typical examples are time segments obtained by performing a segmentation of the time series in order to discover the most significant segments under some suitable criteria. For instance, suppose that the user is interested in finding time segments in which the price is *constant*. These segments and their representation cannot be provided without looking at the data, and hence must be generated by the GLiDTS system during the extraction task.

Let us remark that in the example instance "Most times between 12 and 15 h the price is constant", it is not evident whether the time interval has been defined a priori (because, for instance, target users are interested in such interval), or obtained a posteriori (because the system has discovered that such intervals match the semantics of the instance).

Deeply linked to protoforms and components is how to assess the *quality* of a protoform's instance, i.e., to which degree the instance must be part of the final collection of messages to be transmitted to the target user. As we shall see, quality has several dimensions, the compatibility of instances and data being an important but small part. We shall discuss the fundamental issue of quality assessment via a quality framework in Section 5.

We introduce the following definitions:

Definition 1. The instance space is the set of all protoform instances that can be built using the knowledge representation formalism.

Definition 2. The semantic space is the power set of the instance space.

Elements of the semantic space represent the semantics of every possible linguistic description. The instance space is comprised of:

- Instances of the lower abstraction level, i.e., those whose quality assessment can be obtained directly from the data, like instances of the protoforms *X is A* and *Q D are A*. These are built using the protoforms in the formalism and the set of components defined either a priori and/or a posteriori.
- Protoform instances with higher abstraction levels, obtained from low-level instances by inference on the basis of expert knowledge. The quality assessment of these instances is usually provided by the inference process on the basis of the quality of protoforms of lower levels. We shall see examples later.

In the following we provide a study of protoforms that have been employed by different authors. Most proposals use a single protoform as formalism, but a few of them consider several protoforms at a time. Let us remark that, since there are different views of the tasks and architecture of GLiDTS systems, in many proposals the authors neither make reference to the notion of protoforms, nor even consider they are using a knowledge representation formalism in some cases. Hence, our methodology for this study has been to interpret the proposals in terms of our view.

3.1.1. Low-level protoforms

The protoform *X is A* has been employed by many authors, for instance:

- In [35] A is a categorical value or a crisp interval of numerical values, ordered in time.
- In [36] this protoform is employed in order to represent change in value in a priori time segments, as well as to represent local features corresponding to exceptions to the global trend.
- In [37], *X* makes reference to the motion of an object in a bidimensional space, and A is a label describing the motion.
- In [38] *X* represent a trend or the sign of a trend, whilst A is a label obtained by combining basic linguistic labels and linguistic hedges.
- In [39] this protoform is employed for representing value, dynamics of change, and convexity of the time series.

Different kind of rules can be employed as protoforms, including fuzzy rules *If X is A, then Y is B* [40,41]. In [39] rules associate the above mentioned value, dynamics of change and convexity. Sequential patterns representing sequences of instances of *X is A* have been employed for approximating a time series in [42,43], and in order to represent sequences of actions performed by a person on the basis of time series of data obtained from sensors [44].

Quantified sentences are one of the most employed protoforms in GLiDTS systems. The protoform *Q D are A* has been used for instance in the following proposals:

- In [45–50] quantified sentences express the amount of time instants in a time segment having a certain value. Time segments, organized as a hierarchy, and linguistic labels about value are defined a priori, in some cases in a multidimensional database structure.
- In [51,49], protoforms about tendency (dynamics of change, duration, and variability) of the series are represented by quantified sentences, which are later employed by other authors, e.g. [52].
- In [53,48] quantified sentences are used in which D and A are characteristics of two different time series that are related via quantification in the same time segment.
- Quantified sentences have been employed for comparison of time series. In [54], comparison is performed on the basis of local change of values as differences between consecutive points. In [55], a time series calculated as the difference in value of the original series is described. Similarity Quantified Expressions are employed in [56] using a kind of fuzzy *generalized quantifier* [57,58] called *temporal similarity semi-fuzzy quantifier*. Comparison in terms of tendency using quantification is proposed in [59].
- In [60] quantified sentences are employed for quantifying the amount of points having smaller values than any previous point and bigger value than any following point from some specified neighborhood of the point.
- Fuzzy quantification based on type-II fuzzy sets are employed in [61,62].

Several authors have provided new protoforms for specific applications, for instance:

- The proposal in [63] uses protoforms obtained as the functional composition of quantified sentences.
- In [64,65], descriptions based on periodicity are provided on the basis of the protoform *M every p unit, X are A*, where M is an adverb (roughly, exactly, etc.), p is the length in time instants, and A is a linguistic label about the value of the series.
- In [66,67] the protoform *Xc is Si in Pk for Tj* represents that person (X_c) is in a given state (S) in a location (P_k) for a time interval of duration T_j. It is applied to information obtained by video analysis.
- In [68] the protoform (quality of) *[X] is [A] with [mse] in a [time window]* is employed, where X represents a behavior, A represents the dynamics of change in a given time segment [time window], measured as in [51], with [mse] a linguistic label about the stability of the change.

3.1.2. *Higher-level protoforms*

- In [11,69], the authors use temporal links and causal relationships between events and/or components corresponding to patterns observed in a collection of time series. They are obtained by inference processes based on expert knowledge and ontologies. The final knowledge is organized in a structure called *tree of related events*.
- In [39] the temporal logic proposed by Allen is employed in order to obtain relations between time intervals when expressing rules.
- In [53,48], quantified sentences obtained from a time series are clustered according to a measure of similarity between sentences proposed by the authors, an each cluster is represented by a medoid, which is a sentence that occupies a central position in the cluster.
- In [21], the different elements and tasks of a GLiDTS system are represented via a hierarchically organized model called *Granular Linguistic Model of a Phenomenon*. The nodes of the hierarchy represent *perceptions*, each one being expressed by means of an associated protoform. There is no restriction as to which protoforms can be employed; in fact, the (many) applications in which this model has been applied (see Section 6) use a wide variety of protoforms, even in the same application. A directed link between perceptions P_1 and P_2 in the hierarchy indicates that perception P_1 *explain* P_2, meaning that protoform instances of a certain perception are calculated by inference on the basis of instances of the perceptions which explain it, using generally fuzzy rules and Mamdani-based inference, but allowing for other kinds of inference. Perceptions which are explained by the input data are called *first order perceptions*, and are usually associated to low-level protoforms like *X is A* or quantified sentences. Perceptions explained by other perceptions are called *second order perceptions*. The model

allows the representation of uncertainty about the perceptions, as well as associating specific linguistic expression mechanisms to each perception.

3.2. Extraction technique

The extraction task takes as input the time series data, that may be comprised of a collection of time series, and provides as output the collection of messages expressing the semantics of the text to be generated, using a predefined knowledge representation formalism. As pointed out previously, this can be seen as a KDD process.

We can distinguish two kind of operations that may be employed in the extraction task:

- **Generating new data series** by using data preprocessing, numerical data analysis, and inference tasks. This step basically consists in obtaining one or several time series from the input ones, with values defined either on the same domain that the originals, or different ones. The new series can be obtained by applying one, or a combination of the following operations, among others:
  – Data transformation (normalization of the value domain of time series data, smoothing, noise elimination and trend estimation, time shift, etc.).
  – Obtaining a time series from two or more series, where the value of the new series in a time instant is obtained from the values of other series in the same instant as the difference between values of two series (typical in some approaches to generating linguistic descriptions of the difference between two series), or by aggregating the values of the original series using some aggregation operator, or by determining a new series by inference using a knowledge base with expert rules, or a neural network, etc.
  – Obtaining a time series from a single input series by calculating the difference between values of consecutive time instants.
  – Obtaining a time series taking values in a set of states using a finite state machine fed by the original time series data. In each time instant, the new series takes as value the state of the machine after being fed with the value of the original series in that instant (that is, the state of the machine after taking as input an initial state and the sequence of values of the series from the first instant to the considered one).
  – Using some time series forecasting predictive technique for generating the expected values of a time series in the future from the historical data.
- **Searching the semantic space** for discovering the set of instances representing the final messages. The discovery process is guided by quality measures for individual instances as well as for sets of instances, that are defined as part of the quality framework, as will be discussed in Section 5.

The computational complexity of the GLiDTS process can be given by any of the two operations. Since the size of the semantic space is exponential on the number of instances, computational complexity of the search task is exponential in principle, and search optimization techniques have to be provided for bounding the search. However, many systems avoid this problem by considering that the choice of a message is independent from the choice of any other, so in practice the search is linear on the number of instances. A particular case is that of systems which simply provide the list of all instances with their corresponding quality measures. When the instance space is reasonably small, no search optimization technique is necessary.

In the next sections we provide an analysis of extraction approaches in the literature. Again, what we show here is the interpretation in terms of our particular view of the GLiDTS process.

3.2.1. Exhaustive search

As pointed out before, exhaustive search is a feasible choice when the instance space is small and/or there is independence between the choice of instances. This kind of search is employed in [70,36] using a priori defined components, in order to provide the best result. Other proposals based on exhaustive search are:

- Refs. [42,43] determine a posteriori components related to the form of the series by calculating the matching of the series with a collection of predefined forms.
- Refs. [68,44] extract patterns corresponding to sequences of actions, learnt by means of an inductive learning mechanism based on Hidden Markov Models. These sequences represent the main actions of a specific behavior.

A Matching system, which uses Regular Grammars to recognize the user activities through the Behavior Database, is employed in order to discover actions performed during time on the basis of time series data provided by sensors, as well as deviations from usual behavior.
- Ref. [38] uses the F-transform operation defined by the authors [71] in order to obtain the trend of the time series, and then perform an exhaustive search on the series looking for time segments which are uniform in trend. F-transform is also used in [41] as an initial step for a subsequent perception-based logical deduction process [72].
- Ref. [37] searches for the combination of linguistic hedge and linguistic label which provides the best description of the average value in a certain time window.
- Refs. [66,67] use a hierarchical system of fuzzy inference for activity reasoning for generating a time series describing states from sensor information. Instances of protoforms corresponding to time segments in which the state is constant are provided.
- In [73] dynamics of change are predefined patterns that are then employed for obtaining a posteriori time segments. A machine learning technique for learning these dynamics is used instead in [74], using a SOM learned with a lvq-type algorithm.
- Quantified sentences are obtained by exhaustive search, which are later used for generating the most representative [75], or anomalies with respect to a certain pattern [48].

3.2.2. Non-exhaustive search

There are several techniques which bound the search space by performing an iterative process in the hierarchy of instances with respect to the abstraction level, consisting of two steps: i) searching instances in a certain abstraction level, starting from the lower level (searching among the low-level instances), and ii) using the obtained instances for generating the instances of the next level by inference. The system assumes that the best instances in a level are always among those obtained from the best instances of the previous level. Among these techniques, we can mention the following:

- In [69,11], a posteriori components are obtained as patterns via signal analysis, from which causal relations are obtained by means of temporal and logical reasoning based on expert rules and ontologies.
- In [21] the iterative process is clearly stated in the design of the Granular Linguistic Model of a Phenomenon for a specific application. The input data may come from different time series data, including a time series of states coming from a fuzzy finite state machine, following a model proposed by the authors [76].

Other approaches are the following:

- In [46,47], Greedy algorithms are used for exploring the instance space by abstraction levels defined by the granularity of time segments, using for this purpose a predefined hierarchy of fuzzy partitions of the time. The algorithm starts from the higher levels, trying to find a good instance for describing the segment. If no instance is good enough, the algorithm tries to describe the segment by finding good instances describing the segments of lower levels which have non-empty intersection with it. A priori components describing values are employed. Also, a posteriori components describing trends are obtained by using linear segmentation in [77].
- Ref. [45] uses candidate generation using the "a priori" property of frequent itemsets for bounding the search.
- Refs. [64,65] perform segmentation using their own clustering method based on watershed with a previous smooth of the time series. Instances are obtained using a methodology called *Detection of Periodic Events* (DPE).
- In [78,79] a multi-objective memetic algorithm based on NSGA-II with the addition of a number of intelligent mutation operators that apply heuristics to improve solutions is employed for exploring the semantic space.

4. Linguistic expression process

As we have mentioned previously, to complete the production of a linguistic description of the time series, it is necessary to turn the set of messages obtained in the knowledge extraction task into a text that suitably communicates the knowledge to the target user.

The choice of the words that make this text clearly depends on both the preferences of the writer (i.e. the system designer) and the linguistic context [35]. However, systems must not ignore the communicative needs of the receiver

Table 1
Some examples of string patterns.

String pattern	Instance
On Q of y?s the resident had P.	On most of the nights the resident had a medium level of restlessness.
On Q of y?s, when the resident had R, he had also P.	On most of the nights, when the resident had a medium level of restlessness, he had also a low level of motion.
ET on Q of y?s the resident had P.	Recently on most of the nights the resident had a medium level of restlessness.
ET on Q of y?s, when the resident had R, he had also P.	Recently on most of the nights, when the resident had a medium level of restlessness, he had also a low level of motion.

(i.e. the target user) [17]. In fact, as the variety of target users increase, the modeling of these communicative needs becomes more difficult, even more if we take into account that each user's needs also vary according to her/his context [80].

Two main kinds of approaches are usually considered in the literature to accomplish this task, namely, template-based approaches and *real* (or NLG standard) approaches [81,82]. Approaches of the first kind produce the text by *simply* manipulating strings, while approaches of the second kind perform a more sophisticated linguistic expression of the knowledge taking into account sentence planning and syntactic issues. Though these two types of approaches are commonly accepted, the classification is not binary and more and more approaches could be considered as *hybrid*.

Most GLiDTS systems that can be found in the literature could be considered mainly template-based because they use in the linguistic expression task little (if any) linguistic knowledge. The output of these systems is one or more instances obtained by performing simple operations (usually substitution) on one or more pre-defined string patterns that suitably transmit the knowledge. This is mostly due to the fact that most approaches use knowledge representation mechanisms based on few protoforms.

As an example, Wilbik et al. in [53] propose a set of string patterns that can be easily obtained from a knowledge representation formalism based on quantified sentences and that can be directly instantiated (see Table 1).

An example of more complex output text can be found in [83]. Fig. 2 depicts some details regarding the *template* (in the authors' words) used in this approach to provide information concerning the use of a driving simulation tool. The instantiation of this template is performed by means of *a relevancy analysis in order to select and compile the relevant sentences into one document highlighting the interesting characteristics of a simulation*. These authors have also used the GLMP approach with *templates* that change the structure according to the validity degrees of the generated sentences, though not in the context of time series description [84,85].

Table 2 contains some examples of the output text generated by some remarkable systems. As can be observed in the table, the complexity of the string patterns varies among systems. Simplest ones produce the text as a direct translation of each element of the knowledge representation model, while more complex ones use more sophisticated linguistic expressions which merge different elements of the knowledge representation model.

In most of the approaches cited in Table 2 the translation from the extracted knowledge to the output text is not complex and direct, though some approaches incorporate some linguistic expression enhancement: for example, Kobayashi and Okumura [42] use a dictionary of verbal expressions in order to generate the text by means of rules defined on previously defined patterns that have been detected in the series and Castillo et al. [95] perform some aggregation on the generated sentences in order to reduce the length of the generated text.

Beyond these GLiDTS systems, we can find some others that better deal with linguistics and domain knowledge when building the output text. These systems pay more attention to matters such as morphology, syntax and punctuation. As an example of these GLiDTS systems, Table 3 contains some instances of the generated text together with some remarks regarding these approaches.

Some of the remarks presented in the table are relative to issues that can be considered both concerning knowledge extraction and linguistic expression. This is the case, for example, of the temporal abstraction carried out in [11,69]. The inference which produce relations between events detected in data analysis is tightly related to language expression in this approach. However, this inference process produces knowledge (e.g. temporal and causal relations between events) and, thus, should be also considered as part of the knowledge extraction task. The fact which is behind this difficulty of placing issues like the mentioned one in the knowledge extraction task or in the linguistic expression

i General observations: this section holds various subsections to describe in detail the different aspects happened on the simulation exercise.
  a Interaction activity: this subsection gives information about the interaction activity with the analyzed onboard system during the simulation. It states the number of independent interactions and the accumulated time of interaction. A figure indicating the periods of interaction is also included.
  b Comparison of interaction vs. non-interaction: the quantified sentences obtained from y_{i5}^2 and y_{i6}^2 [a] describing the quality of driving are compared trying to highlight the similarities and differences between them. The following sentences show an example:
  "Both during interaction and non-interaction, Most of the time the Driving Quality has been High".
  "During interaction, Sometimes the Driving Quality has been Low, while during non-interaction A few times the Driving Quality has been Low."
  Note that the template provides two different linguistic expressions depending on the results. This subsection also provides a relation of distraction events during the simulation indicating the number of events, and for each event, occurrence time and cause of event, e.g.,
  "During interaction 1 distraction event happened: Event (1) at 100 s: The Vehicle Linearity is Low".
  c Detailed description of events during interaction: in this subsection, the route-cause of events during the interaction are explained individually. Depending on the rules defined in each aggregation function, the GLMP navigates backwards on its branches solving the inverse problem, in order to deduce which rules have been triggered and by which parameters. Each event has an associated time, and so the generation of statements at the exact time of the event is a straight forward task (we identify the antecedents of the triggered fuzzy rules), e.g., *"Event (1) approximate time at 100 s.*
  The Vehicle Linearity is Low:
  The Vehicle Speed is Medium.
  The Percentage of accelerator usage is High.
  The Vehicle is Not Overtaking.
  The Lateral Position is Rapidly Decreasing."
  Each event description is also accompanied with video frames and figures showing the related input parameters.
ii Conclusion: this section provides the estimated IPD during the simulation exercise.

[a] y_{i5}^2 and y_{i6}^2 are Computational Perceptions in the used GNLP structure.

Fig. 2. Sections of the template of the report used in [83].

task is that natural language is a knowledge representation formalism itself and, therefore, it can be used to infer more sophisticated messages with higher abstraction levels as pointed out in Section 3.

5. Quality framework

The purpose of a GLiDTS system is to provide a target user with an output text which meets his/her necessities. Since in general the semantic space is huge, and the expression language provides a large amount of different ways to express the knowledge obtained from the time series data in the extraction task, the number of possible texts that can be generated by the system is overwhelming. Hence, a formal framework for determining whether one text is appropriate as output for a specific target user, given a certain time series as input, is indispensable. This is what we call *quality framework*, and the suitability for covering the target user's needs is called *quality* of the linguistic description.

Defining the quality framework for a GLiDTS system is a very complex task, specially for general purpose systems, since [18]:

Table 2
Examples of the output text generated by GLiDTS systems with less sophisticated linguistic expression.

Approach	Example instantiated strings
Almeida et al. [63]	Differences: Few alive patients have a low value of HR half of the time, while deceased patients do not. Half of the alive patients have a medium value of HR, very few times, while deceased patients do not. Very few deceased patients have a very low value of HR, half of the time, while alive patients do not. Almost all deceased patients have a very high value of HR, very few times, while alive patients do not. Global: Very few patients have a very low value of HR half of the time. Few patients have a medium value of HR most times.
Alvarez-Alvarez and Triviño [86]	The gait quality of person2 is very high because the gait symmetry is high and the gait homogeneity is high.
Alvarez-Alvarez et al. [87]	At 10:55, the traffic density is extremely high and it is decreasing. The level of service is keeping constant in level F, forced or breakdown flow, queues form behind breakdown points and demand is greater than capacity. The road speed is low and it is keeping constant. At 16:15, the traffic density is medium and it is decreasing. The level of service has changed to level C, stable flow, speeds at or near free-flow and queues may form. The road speed is low and it is keeping constant. At 4:10, the traffic density is extremely low and it is increasing. The level of service is changing from level A to level B, reasonably free flow, the ability to maneuver is only slightly restricted and the effects of minor incidents still are easily absorbed. The road speed is high and it is decreasing. At 4:45, the traffic density is extremely low and it is keeping constant. The level of service has returned to level A, free-flow operation. The road speed is medium and it is increasing. In the morning (from 7:00 to 10:00), the traffic density has been extremely high. The level of service has never been A and B; few times D; sometimes C and E; and many times F. The road speed has been low. There were 4 vehicles speeding. In the afternoon (from 13:00 to 22:00), the traffic density has been low. The level of service has never been A and B; and sometimes C, D, E and F. The road speed has been low. There were not vehicles speeding. At night (from 22:00 to 7:00), the traffic density has been extremely low. The level of service has never been E and F; few times C and D; sometimes B; and many times A. The road speed has been medium. There were many vehicles speeding. During the whole day, the traffic density has been medium. The level of service has been few times B, D, E and F; and sometimes A and C. The road speed has been low. There were many vehicles speeding.
Anderson et al. [88]	Derek is walking in the lab for a moderate time. Derek is in-between in the lab for a short time. Derek is walking in the lab for a short time. Derek is motionless-on-the-chair in the lab for a moderate time.
Castillo-Ortega et al. [47]	At least 70% of the days with mild weather, the patient inflow is high or medium. Most of the days with cold weather, patient inflow is low or very low. Most of the days of May, patient inflow is medium. Most of the days of June, patient inflow is medium. Most of the days of July, patient inflow is high. At least 70% of the days of August, patient inflow is very high. Most of the days of September, patient inflow is high.
Castillo-Ortega el al. [46]	Most days of the second lustrum, both series exhibit a local change with the same sign. At least 70% of the days of second half of year 2000, both series exhibit a local change with the same sign. Most days of year 2001, both series exhibit a local change with the same sign. At least 70% of the days of year 2003, both series exhibit a local change with the same sign. Most days of year 2004, both series exhibit a local change with the same sign.
Castillo-Ortega et al. [55]	At least 70% of the time with cold weather, the patient inflow is much higher in center A than in center B. At least 70% of the time with hot weather, the patient inflow is lower or much lower in center A than in center B. At least 70% of the time with cold to hot weather, the patient inflow is higher or similar in center A than in center B. The patient inflow difference between centers A and B in September presents variability. At least 70% of the time in October, the patient inflow is similar or lower in center A than in center B. Most of the time in November, the patient inflow is much higher or higher in center A than in center B.

(*continued on next page*)

Table 2 (*continued*)

Approach	Example instantiated strings
Delgado et al. [89]	It seems you are going to: To go out home. and you have forgotten: MobilePhone, Keys.
Díaz-Hermida and Bugarín [50]	In most of years preceding 1995, increasings in oil production were negative or slightly positive.
Díaz-Hermida et al. [56]	Rainy days and moderately cold days were associated in January. High values of the asset market and small interest values were associated in the nineties. The correlation between the risky high pressures and the high temperatures has been very high in the last few seconds.
González-Villanueva et al. [90]	The quality of the Sun Salutation execution of subject 5 is medium because the symmetry is high, the stability is high and the rhythm is variable.
Kacprzyk and Wilbik [91,59]	*Most* of the long, simultaneous segments are similar. In most cases if the dynamics of change of the WIG index was constant, dynamics of change of the fund was indifferent.
Kacprzyk et al. [51]	Most of trends are of a large variability. Most of slowly decreasing trends are of a large variability. Trends that took most of the time are of a large variability. Slowly decreasing trends that took most of the time are of a large variability.
Kobayashi and Okumura [42]	At the morning session, sell order was ahead. Afterwards, trading was steady, therefore, the width of rising was small. At the afternoon session, trading was continuously rising. Therefore, The width of rising was expanded. At the closing session, the prices were decline.
Kobayashi et al. [43]	time: 5 Someone comes from the right side. time: 9 Someone sits on the chair. time: 22 Someone stands up. time: 30 Someone sits on the left chair.
Laurent [45]	Most middle sales in 2014 are in January.
Moyse et al. [64].	Approximately every 20 hours, the data take high values.
Novák et al. [38]	Trend of the whole series is stagnating. Slot 1 (time 23–32): clear decrease, Slot 2 (time 70–127): negligible decrease, Slot 3 (time 92–115): small increase, Slot 4 (time 116–127): fairly large decrease.
Ros et al. [68]	Behavior is constant with low Instability in a time period equal to 30 days.
Sánchez-Valdés and Triviño [92]	Currently, the signal is in the state q1. The amplitude of the signal is positive, it is increasing, and the duration is normal to stay in q1 and too short to change to q2. Currently, the model cannot explain the situation. The amplitude of the signal is positive, it is staying, and the duration is too long to stay in q1 and normal to change to q2. Currently, the model cannot explain the situation. The amplitude of the signal is zero, it is decreasing, and the duration is normal to stay in q1 and too short to change to q2. Currently, the model cannot explain the situation. The amplitude of the signal is very negative, it is decreasing, and the duration is normal to stay in q3 and too short to change to q1.
Sánchez-Valdés et al. [84]	During these 40 minutes you walked for a short time and you were sitting most of the time. You have burned very few calories. In order to meet your objectives, you should increase your activity level. During this week you have consumed less energy than it was expected. The energy consumption has been very similar every day of the week. The day of the week you have walked less time was on Thursday, however, on Sunday you have been walking far longer than the other days. On Tuesday and Saturday, you have spent much more time standing than other days. In addition, you spend too much time sitting, so you should reduce as much as possible the time spent sitting.
Sjahputera et al. [37]	A is moving mostly to the left but somewhat upward.
Stepnicka et al. [41]	If the number of cars sold in the current year is more or less small and the biannual sales increment is negative with more or less medium strength, then the upcoming biannual increment will be negative and have medium strength.

Table 2 (*continued*)

Approach	Example instantiated strings
Triviño and Van der Heide [76]	1. First a Short Straight 2. Then a Medium Left Curve 3. Then a Long Straight 4. Then a Soft Right Curve
Triviño et al. [93]	The occupancy in the north ROI (Region of Interest) is high. The movement in the north ROI is high. The number of vehicles going through the north ROI is high. Usually, the roundabout is medium filled.
Umano et al. [36].	It is globally approximately constant and there are a little smaller and a little larger points in the third term.
Van der Heide and Triviño [94]	ABOUT TWO THIRDS of the days the consumption in the MORNINGs is LOWER THAN the consumption in the AFTERNOONs. MOST of the days the consumption in the MORNINGs is lower than the consumption in the EVENINGs. ABOUT TWO THIRDS of the days the consumption in the MIDDAYs is LOWER THAN the consumption in the EVENINGs. ABOUT TWO THIRDS of the days the consumption in the AFTERNOONs is LOWER THAN the consumption in the EVENINGs. MOST of the days the consumption in the AFTERNOONs is HIGHER THAN the consumption in the NIGHTs. ABOUT TWO THIRDS of the days your consumption is LOW during the MORNINGs. MOST of the days your consumption is HIGH during the EVENINGs.
Wilbik and Kaymak [60]	The Resident is getting less active.

- An important part of quality assessment is clearly subjective and context dependent, as the linguistic description is intended to satisfy the needs of a certain user in a given context.
- Quality of linguistic descriptions has many different facets, that we call *quality dimensions*, so in order to assess quality we have to identify and to assess all its dimensions [19]. However, there is no agreement about neither the set of dimensions to consider (which may depend on the specific system), nor the way to evaluate them, either quantitatively or qualitatively [18].
- It is very usual to find a strong interrelation between different quality dimensions, with the important problem that some of them are opposite, contradictory, and/or have a negative correlation. This means that there is in general no such thing as the *optimum* or *best* linguistic description for a given time series data. On the contrary, GLiDTS systems afford a multiobjective optimization problem, in which there are several conflicting objectives.

The dimensions of the quality framework can be related to the instances in the instance space, to sets of instances, as well as to different aspects of the linguistic expression. Hence, the quality framework affects the whole GLiDTS process, guiding the knowledge extraction process, the linguistic expression process, and the final validation of the results.

The quality framework is not comprised of dimensions only; for each dimension we need to provide a specific measure, or some other mechanism, allowing us to determine whether one linguistic description is incomparable, indistinguishable, or better than another with respect to a certain dimension. That is, associated to each dimension, we need to provide a *preference relation* (i.e., a reflexive and transitive binary relation – a preorder) on the domain of linguistic descriptions [18] (ideally, a total order). This is not easy, and in many systems, some dimensions are computationally represented, whilst for others they rely on the target user. In such cases, for the system to use that dimension in the generation process, it is necessary that the user provides directly the preference relation, or to use some machine learning mechanism for obtaining it.

Though not a quality framework in the sense that it defines no specific preference relations, the maxims of Grice [97], which are a well-known set of pragmatic rules for linguistic communication, have inspired the dimensions considered by several quality frameworks for GLiDTS systems in the literature [15,83,49]. The maxims of Grice are the following:

Table 3
Examples of text generated by GLiDTS systems that pay more attention to linguistics.

Approach	Example of generated text	Remark
SUMTIME-MOUSAM (Reiter et al. [35])	W 8″13 backing SW by mid afternoon and S 10″15 by midnight.	In the linguistic expression, the system uses *weatherese* (i.e., weather sublanguage) instead of conventional English, by building special grammar rules based on an analysis of human-written forecasts. The machine learning algorithm C4.5 is used to learn classifiers to predict what time phrase would be used.
BT-45 (Portet et al. [69])	You saw the baby between 14:10 and 14:50. Heart Rate (HR) = 159. Core Temperature (T1) = 37.7. Peripheral Temperature (T2) = 34.3. Transcutaneous Oxygen (TcPO2) = 5.8. Transcutaneous CO2 (TcPCO2) = 8.5. Oxygen Saturation (SaO2) = 89. Over the next 30 minutes T1 gradually increased to 37.3. By 14:27 there had been 2 successive desaturations down to 56. As a result, Fraction of Inspired Oxygen (FIO2) was set to 45%. Over the next 20 minutes T2 decreased to 32.9. A heel prick was taken. Previously the spo2 sensor had been re-sited. At 14:31 FIO2 was lowered to 25%. Previously TcPO2 had decreased to 8.4. Over the next 20 minutes HR decreased to 153. By 14:40 there had been 2 successive desaturations down to 68. Previously FIO2 had been raised to 32%. TcPO2 decreased to 5.0. T2 had suddenly increased to 33.9. Previously the spo2 sensor had been re-sited. The temperature sensor was re-sited.	The authors performs *temporal abstraction* to group sequences of events that have been formed as part of data abstraction into a single event frame. The system establishes links (e.g. causal and part-whole relations). A discourse manager handles expression of time and temporal relations (which are expressed using adverbials and tenses).
Portet and Gatt [11]	1. He is currently on nasal CPAP in air, having been extubated today [...] Prior to extubation his SpO2 and HR showed compromise during handling with desaturations and HR decelerations. 2. The baby was moved from SIMV to CPAP. He was extubated and underwent oral suction. This must have caused the instability in HR and SpO2.	Temporal relations are dealt with by means of a fuzzy approach. The paper discusses how abstracted temporal relations, and other related relations (e.g. causality), can be exploited to communicate uncertainty using expressions such as epistemic modals.
GALIWeather (Ramos-Soto et al. [96])	There will be clear skies at the beginning and towards the middle of the term, although at the end they will be very cloudy. We expect precipitations on Thursday morning. The temperatures will be normal for the minimums and high for the maximums for this period of the year, with minimums in notable increase and maximums without changes.	The authors define their system as hybrid: *On one hand, we have defined templates in structured text files which contain generic natural language sentences for the simpler variables (cloud coverage, temperatures and wind). On the other hand, we have designed and implemented the generation of natural language sentences for precipitation inspired by standard NLG methodologies.* The systems includes a module to express precipitation which addresses redundancy or length excess in the obtained descriptions.

- Maxim of Quality[4]: Try to make your contribution one that is true. More specifically:
  1. Do not say what you believe to be false.
  2. Do not say that for which you lack adequate evidence.
- Maxim of Quantity:

[4] Let us remark that *quality* here is employed in a narrower sense that we employ when we speak of *quality framework*. Dimensions and relations related to this maxim are just one part of the frameworks employed in the literature.

   1. Make your contribution as informative as is required (for the current purposes of the exchange).
   2. Do not make your contribution more informative than is required.
- Maxim of Relevance: Be relevant.
- Maxim of Manner: Be perspicuous. More specifically:
   1. Avoid obscurity of expression.
   2. Avoid ambiguity.
   3. Be brief.
   4. Be orderly.

We are going to use the Gricean maxims as a basis for organizing our study of the quality framework in the following sections. Notice that we do not claim that all dimensions of a quality framework are expected to fit into these maxims (though existing quality frameworks in the literature fit well, as we shall see). We have found them simply to be a convenient way, under certain interpretations, for structuring our discussion.

5.1. Maxim of quality

This maxim has to be taken into account primarily in relation to the knowledge representation formalism. The first specific maxim can be interpreted in our scheme as: *low-level instances have to accurately match the data, and higher-level instances have to be obtained via sound inference techniques*. The second one, as: *protoforms able to represent the uncertainty of the knowledge must be employed when necessary*; this is the case for instance when providing a linguistic description of a time series obtained by forecasting models which provide some measure of the uncertainty of the prediction, both participating in the final linguistic expression (e.g. *It is likely* that it will rain tomorrow).

The first specific maxim affects several aspects of the extraction task, particularly:

- The calculation of *a posteriori* instance components by analyzing the time series, where the degree of matching between data and components must be assessed. This is computationally represented by measures of goodness of clusters, average error in the approximation of time series using linear segmentation, measures of compatibility of pattern instances with data provided by pattern recognition models, etc.
- Assessment of protoform instances by measuring the degree to which the knowledge expressed by the instance is an *accurate* description of the data. A procedure for measuring the degree of matching between instances and data is a mandatory part in the definition of low-level protoforms. For higher level protoforms, the inference technique must provide instances which are logical consequences (for suitable logic and consequence operator) of instances of low-level protoforms, together with the accuracy of the higher level instances, calculated on the basis of that of low-level instances.

With respect to the most employed protoforms, the assessment of the accuracy of instances of *X is A* is very simple, and given by the membership of the actual value of *X* to the fuzzy set *A*, but this is not the case with quantified sentences, for which there are many different approaches. GLiDTS systems have employed Zadeh's approach based on compatibility of a quantifier with the sigma-count cardinality [63,48], the Sugeno integral [73] and a modified version of it [98], the Choquet integral [99], Yager's method based on OWA [100], method GD of one of the authors in [46,47], and also own methods based on generalized quantification in [70,56,101,50,49]. A review of techniques for the evaluation of quantified sentences including these and other methods can be found in [102].

5.2. Maxim of quantity

This maxim concerns mainly the knowledge representation, and can be interpreted in several ways. On the one hand, it can be closely related to the *level of abstraction* employed in describing the data. Given a time series data, we could linguistically describe it by means of a collection of sentences of a very low level of abstraction of the form *The value of the series in time t is v*; this would be a most accurate and precise expression, and totally informative about the series, but one that will be very complex for the target user to interpret. On the other end of the scale we could find a description like *The values of the series are between v and w*, which is very easy to understand, but very little informative. Higher levels of abstraction correspond to less precise but more easily understandable (less complex) descriptions. Considered exclusively from this point of view, the process of linguistic description of data can be seen as finding the most suitable level of abstraction for presenting knowledge about the data to the target user, see also [16].

Going beyond this aspect of a linguistic description as a precision-complexity tradeoff, this maxim can also be interpreted in terms of *usefulness*, particularly when the objective is to provide knowledge which is useful for decision making. As pointed out by van Deemter, *utility should play a crucial role in our thinking about language generation and production* [9].

Techniques that have considered measures relevant to this maxim are [103], where the *specificity* (linked to the granularity) is measured via a so-called *degree of fuzziness*, [104,105], where different measures of informativeness are used, and [106], where a degree of *appropriateness* representing the novelty is proposed.

5.3. Maxim of relevance

This maxim concerns again mainly the knowledge representation, and points out that the linguistic description has to be relevant to the target user, in the sense of matching his/her final objectives. Descriptions which match perfectly the data and are understandable may not be relevant to the user if they describe aspects of the data that are out of his/her interests and needs. Also, descriptions may be very informative in the sense of being useful for decision making in a decision problem which is of no interest to the user. These are the situations that are to be avoided following the maxim of relevance.

For achieving this, *systems that generate textual summaries should be controlled by models of the particular interests and tasks of specific users* [15]. In many GLiDTS systems, the model of the relevant aspects of the data is prefixed and incorporated to the system's design, so that only knowledge which is relevant to the user is provided. However, there are GLiDTS systems which provide linguistic descriptions as answer to a user's query, in which relevance of the knowledge to the request has to be assessed during the generation process, so more general models are needed.

One measure related to relevance is the support, widely employed in KDD, which have been employed also in GLiDTS quality frameworks [107]. It has also been evaluated by means of target users [19]. For instance in [105] the relevance of a text was defined in terms of whether the events it mentioned have some relationship to the clinical actions which are appropriate for that scenario. In [3] the evaluation is performed by asking users to make decisions based on the generated texts, and measuring the quality of these decisions. User-based evaluation of relevance has been employed as well in [49,83].

5.4. Maxim of manner

This maxim is mainly concerned with the linguistic expression, though aspects like ambiguity and above all brevity are also to be taken into account in the extraction task. In [15] the following interpretation of this maxim for the linguistic description of time series is provided:

1. Avoid obscurity of expression – information should be expressed in the most appropriate linguistic manner.
2. Avoid ambiguity – summary texts should give a consistent message about what is happening.
3. Be brief – summarize only the important aspects of the input data, use aggregation and ellipsis to reduce length.
4. Be orderly – describe events using a consistent ordering strategy.

We agree with this interpretation. Let us point out some remarks about each of them:

1. Expressing information in the most appropriate linguistic manner can be discussed from different points of view. However, a key point is that of the pragmatic competence of the target user, i.e., the expression must be suited to the user. In [17], the importance of incorporating user's preferences in the expression task is stressed. In order to guarantee this, supervised Machine Learning techniques have been employed for learning how to perform linguistic expression subtasks such as lexicalization, etc. using user evaluation on corpus of linguistic descriptions [16].
2. Avoiding ambiguity has also multiple facets. A widespread technique in GLiDTS systems related to this issue is that of choosing the instance of a protoform with higher accuracy for describing a certain feature, discarding the

rest, and avoiding situations like having in the same description *Temperature is high* and *Temperature is medium*. This is also useful for bounding the semantic space. This aspect is also related to sound inference techniques for generating high level instances, which avoids contradictions in the obtained messages.

3. Brevity is a very important aspect of linguistic descriptions. Measures of brevity of the linguistic description have been employed for instance in [46,78,18].
4. Order is part of the linguistic expression tasks known as *document planning*, and is widely employed in GLiDTS systems. For instance, it is common to begin the linguistic descriptions with messages of high level of abstraction, and then focusing on particular areas or features of interest, or exceptions/anomalies with respect to usual patterns, etc.

5.5. The key role of the target user

We finish this section by stressing again the key role played by the target user in defining the quality framework. Though the framework is to be specified by the designer of the GLiDTS system, the specification has to take into account the target user in terms of his/her objectives, preferences, and needs. Though general measures may be employed for assessing some of the dimensions of the quality framework, the specification of thresholds for accuracy and other measurable aspects of instances and whole linguistic descriptions is affected by the target user.

A more tangible role is played in the final evaluation of the system's output, which can be employed for a dual purpose: first, to evaluate the system itself; second, for learning user's preferences in different aspects of the generated text, as we have seen in the previous section, using machine learning techniques.

The evaluation by means of users is a particular case of what is known as *extrinsic evaluation*, also called *evaluation in use*, in contraposition to *intrinsic evaluation*, which considers an isolated NLP system and characterizes its performance mainly by measuring syntactic similarity with respect to a gold standard result, pre-defined by the evaluators. On the contrary, extrinsic evaluation assess the performance of the system in terms of its utility with respect to the human user [35,69,11,86,83,49].

For instance, systems have been evaluated by asking users to make decisions based on the generated texts, and measuring the quality of these decisions [3]. In [108] the frequency and type of post-edits performed by humans of the generated text is employed as a measure of system's performance.

In [49,83] questionnaires have been used in order to ask experts in linguistic description in a certain domain about the quality of the output text. In [49], in the setting of weather forecasting, the questionnaire contains questions about the accuracy of the content and the correctness of its form, regarding Relevance ("Does the forecast include all the kind of information the expert would include?"), Truthfulness ("Does the included information in the forecast reflect the numeric-symbolic forecast correctly?"), and Manner ("Does the forecast express the information properly?", "Is it well formatted?").

In [83], the questionnaire have several questions for which five possible answers in form of numeric scale of evaluation in [1,5] are possible. Question 1 considers subjective relevance, question 2 considers the intersubjective relevance, questions 3 and 4 deal with evaluating the truthfulness of a report, and questions 5 and 6 evaluate if the report uses the adequate vocabulary, if the order of the ideas is the most appropriate or if the used expressions are the right ones.

6. Applications

As can be observed in the works cited in this paper, the development of GLiDTS systems is growing more and more in the last decade, specially in the field of soft computing practitioners. Among the proposals that we have cited in the previous sections, we can find special purpose applications that have been designed to produce linguistic descriptions of time series data in many specific fields:

- One of the fields that have received more attention is Healthcare. In this field we can find systems like BT-45, which generates textual summaries of sensor data from a neonatal intensive care unit [69]. Almeida et al. [63] also analyze in their work intensive care unit data regarding abdominal septic shock in patients admitted to 70 different hospitals in Germany. There also exist systems aimed at the care of elder people like the proposal of Delgado et al. [89], which generates texts that describe elder people's behavior and produces alarms in dangerous situations, or the proposal of Wilbik et al. [48] also regarding to the monitoring of activities, in this case by describing different events that may occur during the night from sensor data that measure restlessness in bed and bedroom motion of residents of an eldercare facility. Ros et al. [68] also focus on behavior analysis in eldercare and Anderson et al. [66,67] perform activity analysis of elder people by means of linguistic description of video sequences. The linguistic description of the human gait quality has been studied by Alvarez-Alvarez and Triviño [86] and, focusing on quasi-periodicity, by Sánchez-Valdés and Triviño in [85] using the general approach for this feature proposed in [92]. Finally, we can mention the work of González-Villanueva et al. regarding the description of rehabilitation exercises [90].

- *Meteorology* is another topic that has received great attention. For example, SUMTIME-MOUSAM system [35] generates weather forecast texts from numerical weather prediction data, as well as GALIWeather [96], which automatically generates textual short-term weather forecasts too, in this case for every municipality in Galicia (Spain), using the real data provided by the Galician Meteorology Agency (MeteoGalicia).
- Time series are very frequent in stock market and, thus, some of the analyzed proposals focus on this topic. For example, Kacprzyk and Wilbik [109,91] proposal evaluates the performance of an investment fund over a past period.
- Regarding ecology, a GLiDTS prototype capable to provide a linguistic description of ecological phenomena is proposed by Romero and Moreno in [110] and Van der Heide and Triviño [94] focus on the linguistic description of energy consumption data of households that can be send as a complement of the invoices that companies send to clients.
- Other applied works can be found related to the generation of textual summaries of sensor data from a gas turbine (SUMTIME-TURBINE system [111]), to the reporting in driving simulation environments (see [83]), to the description of perceptions while driving (see [76]) or to the linguistic description of traffic (see [93]).

More general approaches can be found related to the verbalization of human behaviors that can be observed in video data (see, for example, [43]) or regarding the descriptions of human activities based on mobile phone's accelerometers (see [84]).

Finally, we can find some proposals that have been applied in business intelligent tools, like OLAP systems, that can be applied in a wide variety of applications domains to assist in the decision making process. Examples of this can be found in [45] and in the system Linguistic F-Cube Factory [112].

7. Conclusions

GLiDTS systems are intelligent systems which design can be very complex, requiring many different skills and expert knowledge (time series analysis, semantic representation for linguistic terms and expressions, with special attention to different types of uncertainty, expert knowledge acquisition and representation, inference techniques, the many aspects of natural language generation with special attention to pragmatics, user preferences and quality measures, search optimization techniques, etc.).

In this paper, we have presented a general architecture for GLiDTS systems, from which we conceive the creation of this type of data-to-text systems and which includes a description model with three main pillars, namely, knowledge representation, linguistic expression and quality assesment. We have carried out a deep analysis of the main tasks of generating linguistic descriptions of time series and the main components of the mentioned description model, putting under a unifying perspective tools and techniques coming from a representative survey of the proposals that can be found in the literature in many applications domains. The analyzed contributions mainly arise from NLG and fuzzy set theory. In these approaches, conventional time series analysis techniques like forecasting, wavelet analysis, segmentation, clustering, and so on, turn out to be useful component in the processing of this type of systems.

In Section 7.1 we discuss about the importance of fuzzy set theory and other uncertainty representation models in GLiDTS systems, including the not yet exploited potential contributions. Finally, in Section 7.2 we provide a final reflection.

7.1. Fuzzy sets in GLiDTS systems

The importance of fuzzy set theory and extensions in the GLiDTS process is visible in the light of our study in previous sections.[5] As we have seen, fuzzy set theory can be used for message representation formalisms, and we have shown in our analysis that formalisms employed in GLiDTS systems can be seen as structured collections of protoforms within the GTU of Zadeh. The necessity of dealing with different kinds of uncertainty in GLiDTS systems is widely acknowledged by the authors working on this area, including those coming from the NLG community, which have employed fuzzy set theory and extensions in their approaches.

As we mentioned in the introduction, fuzzy set theory plays a key role in filling the semantic gap between data and linguistic terms and expressions, and in dealing with the different kinds of uncertainty inherent to the linguistic expression of knowledge about real world data [9–11], in line with the well-known suitability and high potential of fuzzy set theory for representing the semantics of natural language expressions [12]. Let us remark that one of the ideas behind the development of fuzzy set theory was the precision-complexity tradeoff of systems. By gradually diminishing the precision in the representation, we can express knowledge about time series data using higher levels of abstraction, diminishing the complexity of understanding data and increasing its usefulness. Fuzzy set theory is a most suitable tool for that purpose. Of course, we consider within fuzzy sets the use of inference techniques for knowledge affected of uncertainty, that can be employed for deriving linguistic descriptions, and have also a crucial part in the assessment of protoforms.

In addition to the abovementioned contributions, several well-developed research areas within fuzzy set theory can provide significant contributions to the study of GLiDTS systems. A non-exhaustive list of them is the following:

- In our view, the extraction task of GLiDTS systems can be considered a KDD task. In this sense, using KDD patterns and algorithms in GLiDTS systems has an enormous potential that has been explored in hardly a couple of proposals, with well-developed quality frameworks [113]. Also, our analysis has highlighted the widespread use of Machine Learning techniques in different aspects of GLiDTS systems, and the high potential it offers in learning patterns in time series data and representations of several dimensions of the quality framework from corpus of linguistic descriptions assessed by users, particularly those related to the linguistic expression [16]. As we mentioned before, in KDD and Machine Learning, fuzzy set theory has the potential to produce models that are more comprehensible, less complex, and more robust, being especially useful for representing "vague" patterns and modeling and processing various forms of uncertain and incomplete information [8]. Fuzzy patterns and models have been employed in KDD [25,26,29–33]. Also, fuzzy set theory has been shown to be a more suitable approach than probability for learning to rank [114], with applications in learning preference relations [115], which can play a significant contribution for quality frameworks and linguistic tasks like lexical substitutions [116].
- Some problems related to the expression of uncertainty about the obtained data can be solved using protoforms already proposed by Zadeh in his GTU, for instance, Van Deemter speaks of expressing our confidence in the result. Consider for instance the protoform exemplified by the linguistic expression of an instance: *It is very likely that most of the days, the weather is good.*
- Quantification is on the basis of a wide range of protoforms. However, the full potential of state-of-the-art fuzzy quantification has not been employed in GLiDTS systems. Approaches for fuzzy generalized quantification have been proposed by several authors, see [102]. Closely related, recent developments that have improved the representation of fuzzy cardinality and entropy measures are also very promising. Other fuzzy measures and concepts such as probability of fuzzy events and fuzzy probabilities, mixing different kinds of uncertainty, can play an important role.
- For use in the preprocessing task of knowledge extraction, several techniques for time series analysis and forecasting are available in the literature. Let us remark that in this and other fuzzy set theory areas, the word *linguistic* is employed for pointing out the use of linguistic labels, having nothing to do with the generation of natural language. Fuzzy clustering is also very useful for providing a posteriori instance components, as well as fuzzy pattern recognition techniques. The use of fuzzy aggregation operators for generating time series as aggregation

[5] Other Soft Computing techniques have been applied in a few GLiDTS systems, as we have seen (Evolutionary Computation, Neural Networks, Wavelet analysis for preprocessing, etc.) and offer a plethora of techniques with the potential of being used for improving GLiDTS systems.

of others, and for integrating the different dimensions of the quality framework when these are based on fuzzy measures (or appropriately fuzzified [18]), is another interesting future avenue to explore.
- Representation of preference relations are fundamental for the quality framework, in close relation with fuzzy measures, as well as with decision making techniques. In these areas, fuzzy set theory has provided and is still providing many significant contributions.
- Inference techniques for knowledge affected of uncertainty can be employed for deriving linguistic descriptions, in particular, fuzzy description logics and the corresponding ontologies [117], in which several proposals include fuzzy quantification.
- Representing and dealing with time series data coming from uncertain and/or imprecise data.

7.2. Final reflection

The need for general purpose GLiDTS systems should be cited as one of the directions towards to which the research community must focus its efforts. Systems like multidimensional databases or temporal databases can substantially improve their communication with the user through the use of natural language [118]. Though general purpose NLG systems are very complex to develop, the challenge is more affordable if they are restricted to the description of time series data. Some approaches have been developed in the recent years, for example the OLAP systems with linguistic capabilities that we cited in previous sections (see [45,112]), but there is still much to do in order to produce richer linguistic descriptions of time series.

In fact, as the knowledge regarding the series that the systems aims at transmitting grows, the complexity of the two main tasks we have placed in the heart of GLiDTS systems increases; not only more varied knowledge representation formalisms and extraction techniques are needed, but also linguistic expression must be developed in a more sophisticated way, taking advantage of the power of natural language to suitably transmit the knowledge to the target user. This fact, yet envisioned by Reiter in [81], supposes nowadays a call to the joint use of the many KDD and Machine Learning techniques, together with the more distilled methods to produce text in the NLG area, as the way to face the challenge of producing general purpose GLiDTS systems.

Acknowledgements

This work has been partially supported by the Spanish Ministry of Economy and Competitiveness and the European Regional Development Fund (FEDER) under project TIN2014-58227-P “Descripción lingüística de información visual mediante técnicas de minería de datos y computación flexible”.

References

[1] M. Petre, Why looking isn’t always seeing: readership skills and graphical programming, Commun. ACM 38 (6) (1995) 33–44.
[2] E. Reiter, An architecture for data-to-text systems, in: Proceedings of the Eleventh European Workshop on Natural Language Generation, ENLG ’07, Association for Computational Linguistics, Stroudsburg, PA, USA, 2007, pp. 97–104.
[3] E. Reiter, A. Gatt, F. Portet, M. van der Meulen, The importance of narrative and other lessons from an evaluation of an NLG system that summarises clinical data, in: M. White, C. Nakatsu, D. McDonald (Eds.), INLG, The Association for Computer Linguistics, 2008.
[4] P.A. Carpenter, P. Shah, A model of the perceptual and conceptual processes in graph comprehension, J. Exp. Psychol., Appl. 4 (2) (1998) 75–100.
[5] L. Ferres, A. Parush, S. Roberts, G. Lindgaard, Helping people with visual impairments gain access to graphical information through natural language: the igraph system, in: K. Miesenberger, J. Klaus, W. Zagler, A. Karshmer (Eds.), Computers Helping People with Special Needs, in: Lect. Notes Comput. Sci., vol. 4061, Springer, Berlin/Heidelberg, 2006, pp. 1122–1130.
[6] E. Reiter, R. Dale, Building Natural Language Generation Systems, Cambridge University Press, New York, NY, USA, 2000.
[7] G. Piateski, W. Frawley, Knowledge Discovery in Databases, MIT Press, Cambridge, MA, USA, 1991.
[8] E. Hüllermeier, Fuzzy sets in machine learning and data mining, Appl. Soft Comput. 11 (2) (2011) 1493–1505.
[9] K. van Deemter, Utility and language generation: the case of vagueness, J. Philos. Log. 38 (6) (2009) 607–632.
[10] K. van Deemter, Vagueness facilitates search, in: M. Aloni, H. Bastiaanse, T. de Jager, K. Schulz (Eds.), Amsterdam Colloquium on Logic, Language and Meaning, in: Lect. Notes Comput. Sci., vol. 6042, Springer, 2009, pp. 173–182.
[11] F. Portet, A. Gatt, Towards a possibility-theoretic approach to uncertainty in medical data interpretation for text generation, in: D. Riaño, A. ten Teije, S. Miksch, M. Peleg (Eds.), KR4HC, in: Lect. Notes Comput. Sci., vol. 5943, Springer, 2009, pp. 155–168.
[12] L.A. Zadeh, From computing with numbers to computing with words – from manipulation of measurements to manipulation of perceptions, Appl. Math. Comput. Sci. 12 (3) (2002) 307–324.

[13] R.R. Yager, On linguistic summaries of data, in: Knowledge Discovery in Databases, 1991, pp. 347–366.
[14] K. Kukich, Knowledge-based report generation: a technique for automatically generating natural language reports from databases, in: Proceedings of the 6th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '83, ACM, New York, NY, USA, 1983, pp. 246–250.
[15] S. Sripada, E. Reiter, J. Hunter, J. Yu, Generating English summaries of time series data using the Gricean maxims, in: L. Getoor, T.E. Senator, P. Domingos, C. Faloutsos (Eds.), KDD, ACM, 2003, pp. 187–196.
[16] P.K.V. Sowdaboina, S. Chakraborti, S. Sripada, Learning to summarize time series data, in: A.F. Gelbukh (Ed.), CICLing (1), in: Lect. Notes Comput. Sci., vol. 8403, Springer, 2014, pp. 515–528.
[17] K. van Deemter, A. Gatt, I. van der Sluis, R. Power, Generation of referring expressions: assessing the incremental algorithm, Cogn. Sci. 36 (5) (2012) 799–836.
[18] R. Castillo-Ortega, N. Marín, D. Sánchez, A. Tettamanzi, Quality assessment in linguistic summaries of data, in: S. Greco, B. Bouchon-Meunier, G. Coletti, M. Fedrizzi, B. Matarazzo, R.R. Yager (Eds.), IPMU (2), in: Commun. Comput. Inf. Sci., vol. 298, Springer, 2012, pp. 285–294.
[19] I.H. Khan, G. Ritchie, K. van Deemter, The clarity-brevity trade-off in generating referring expressions, in: N. Colineau, C. Paris, S. Wan, R. Dale, A. Belz, R. Dale (Eds.), INLG, The Association for Computer Linguistics, 2006, pp. 89–91.
[20] S. Boyd, Trend: a system for generating intelligent descriptions of time-series data, in: Proceedings of the IEEE International Conference on Intelligent Processing Systems, ICIPS-1998, 1998.
[21] G. Triviño, M. Sugeno, Towards linguistic descriptions of phenomena, Int. J. Approx. Reason. 54 (1) (2013) 22–34.
[22] L. Zadeh, Generalized theory of uncertainty (GTU) – principal concepts and ideas, Comput. Stat. Data Anal. 51 (2006) 15–46.
[23] L.A. Zadeh, A prototype-centered approach to adding deduction capability to search engines—the concept of protoform, in: Annual Meeting of the North American Fuzzy Information Processing Society, Proceedings, NAFIPS 2002, 2002, pp. 523–525.
[24] J. Kacprzyk, S. Zadrozny, Linguistic database summaries and their protoforms: towards natural language based knowledge discovery tools, Inf. Sci. 173 (4) (2005) 281–304.
[25] M. Delgado, N. Marín, D. Sánchez, M. Vila, Fuzzy association rules: general model and applications, IEEE Trans. Fuzzy Syst. 11 (2) (2003) 214–225.
[26] M. Delgado, M. Ruiz, D. Sánchez, J. Serrano, A formal model for mining fuzzy rules using the RL representation theory, Inf. Sci. 181 (2011) 5194–5213.
[27] D. Sánchez, J. Serrano, I. Blanco, M. Martín-Bautista, M. Vila, Using association rules to mine for strong approximate dependencies, Data Min. Knowl. Discov. 16 (3) (2008) 313–348.
[28] F. Berzal, I. Blanco, D. Sánchez, J. Serrano, M. Vila, A definition for fuzzy approximate dependencies, Fuzzy Sets Syst. 149 (1) (2005) 105–129.
[29] E. Hüllermeier, Association rules for expressing gradual dependencies, in: Proceedings PKDD 2002, in: Lect. Notes Comput. Sci., vol. 2431, 2002, pp. 200–211.
[30] H.-W. Koh, E. Hüllermeier, Mining gradual dependencies based on fuzzy rank correlation, in: C. Borgelt, et al. (Eds.), Combining Soft Computing & Statistical Methods, in: AISC, vol. 77, Springer-Verlag, 2010, pp. 379–386.
[31] C. Molina, J. Serrano, D. Sánchez, M. Vila, Mining gradual dependencies with variation strength, Mathw. Soft Comput. 15 (1) (2008) 75–93.
[32] C. Molina, D. Sánchez, J. Serrano, M. Vila, On some new types of fuzzy dependencies, in: Proceedings IPMU 2008, 2008, pp. 1551–1558.
[33] M. Delgado, M. Ruiz, D. Sánchez, New approaches for discovering exception and anomalous rules, Int. J. Uncertain. Fuzziness Knowl.-Based Syst. 19 (2) (2011) 361–399.
[34] L.A. Zadeh, A computational approach to fuzzy quantifiers in natural languages, Comput. Math. Appl. 9 (1) (1983) 149–184.
[35] E. Reiter, S. Sripada, J. Hunter, J. Yu, I. Davy, Choosing words in computer-generated weather forecasts, Artif. Intell. 167 (1–2) (2005) 137–169.
[36] M. Umano, M. Okumura, K. Seta, Improved method for linguistic expression of time series with global trend and local features, in: FUZZ–IEEE, IEEE, 2009, pp. 1169–1174.
[37] O. Sjahputera, J.M. Keller, P. Matsakis, R. Bondugula, Temporal fusion in dynamic linguistic descriptions, in: FUZZ–IEEE, IEEE, 2003, pp. 618–623.
[38] V. Novák, V. Pavliska, I. Perfilieva, M. Stepnicka, F-transform and fuzzy natural logic in time series analysis, in: J. Montero, G. Pasi, D. Ciucci (Eds.), EUSFLAT Conf., Atlantis Press, 2013.
[39] I. Batyrshin, L. Sheremetov, Perception-based approach to time series data mining, Appl. Soft Comput. 8 (3) (2008) 1211–1221, forging the Frontiers – Soft Computing.
[40] V. Novák, M. Stepnicka, J. Kupka, Linguistic descriptions: their structure and applications, in: H.L. Larsen, M.J. Martín-Bautista, M.A. Vila, T. Andreasen, H. Christiansen (Eds.), FQAS, in: Lect. Notes Comput. Sci., vol. 8132, Springer, 2013, pp. 209–220.
[41] M. Stepnicka, A. Dvorák, V. Pavliska, L. Vavrickova, A linguistic approach to time series modeling with the help of f-transform, Fuzzy Sets Syst. 180 (1) (2011) 164–184.
[42] I. Kobayashi, N. Okumura, Verbalizing time-series data: with an example of stock price trends, in: J.P. Carvalho, D. Dubois, U. Kaymak, J.M. da Costa Sousa (Eds.), IFSA/EUSFLAT Conf., 2009, pp. 234–239.
[43] I. Kobayashi, M. Noumi, A. Hiyama, A study on verbalization of human behaviors in a room, in: FUZZ–IEEE, IEEE, 2010, pp. 1–6.
[44] M. Ros, M.P. Cuéllar, M. Delgado, A. Vila, Online recognition of human activities and adaptation to habit changes by means of learning automata and fuzzy temporal windows, Inf. Sci. 220 (2013) 86–101.
[45] A. Laurent, Generating fuzzy summaries from fuzzy multidimensional databases, in: F. Hoffmann, D.J. Hand, N.M. Adams, D.H. Fisher, G. Guimarães (Eds.), IDA, in: Lect. Notes Comput. Sci., vol. 2189, Springer, 2001, pp. 24–33.
[46] R. Castillo-Ortega, N. Marín, D. Sánchez, Linguistic query answering on data cubes with time dimension, Int. J. Intell. Syst. 26 (10) (2011) 1002–1021.

[47] R. Castillo-Ortega, N. Marín, D. Sánchez, A fuzzy approach to the linguistic summarization of time series, J. Mult.-Valued Log. Soft Comput. 17 (2–3) (2011) 157–182.
[48] A. Wilbik, J.M. Keller, J.C. Bezdek, Linguistic prototypes for data from eldercare residents, IEEE Trans. Fuzzy Syst. 22 (1) (2014) 110–123.
[49] A. Ramos-Soto, A. Bugarín, S. Barro, J. Taboada, Automatic generation of textual short-term weather forecasts on real prediction data, in: H.L. Larsen, M.J. Martín-Bautista, M.A. Vila, T. Andreasen, H. Christiansen (Eds.), FQAS, in: Lect. Notes Comput. Sci., vol. 8132, Springer, 2013, pp. 269–280.
[50] F. Díaz-Hermida, A. Bugarín, Semi-fuzzy quantifiers as a tool for building linguistic summaries of data patterns, in: FOCI, IEEE, 2011, pp. 45–52.
[51] J. Kacprzyk, A. Wilbik, S. Zadrozny, An approach to the linguistic summarization of time series using a fuzzy quantifier driven aggregation, Int. J. Intell. Syst. 25 (5) (2010) 411–439.
[52] H. Hatipoglu, F.E. Boran, M. Avci, D. Akay, Linguistic summarization of Europe Brent spot price time series along with the interpretations from the perspective of Turkey, Int. J. Intell. Syst. 29 (10) (2014) 946–970.
[53] A. Wilbik, J.M. Keller, G.L. Alexander, Linguistic summarization of sensor data for eldercare, in: SMC, IEEE, 2011, pp. 2595–2599.
[54] R. Castillo-Ortega, N. Marín, D. Sánchez, Linguistic local change comparison of time series, in: Fuzz–IEEE 2011, 2011.
[55] R. Castillo-Ortega, N. Marín, D. Sánchez, Time series comparison using linguistic fuzzy techniques, in: E. Huellermeier, R. Kruse, F. Hoffmann (Eds.), Computational Intelligence for Knowledge-Based Systems Design, in: LNAI, vol. 6178, Springer-Verlag, 2010, pp. 330–339.
[56] F. Díaz-Hermida, P. Cariñena, A. Bugarín, Efficient evaluation of similarity quantified expressions in the temporal domain, in: J. Lawry, E. Miranda, A. Bugarín, S. Li, M.A. Gil, P. Grzegorzewski, O. Hryniewicz (Eds.), SMPS, in: Adv. Soft Comput., vol. 37, Springer, 2006, pp. 191–198.
[57] J. Barwise, R. Cooper, Generalized quantifiers and natural language, Linguist. Philos. 4 (1981) 159–219.
[58] D. Westerståhl, Branching generalized quantifiers and natural language, in: P. Gärdenfors (Ed.), Generalized Quantifiers, D. Reidel, Dordrecht, 1987, pp. 269–298.
[59] A. Wilbik, J. Kacprzyk, On a benchmark related assessment of the performance of mutual (investment) funds, in: FUZZ–IEEE, IEEE, 2011, pp. 2934–2939.
[60] A. Wilbik, U. Kaymak, Gradual linguistic summaries, in: A. Laurent, O. Strauss, B. Bouchon-Meunier, R.R. Yager (Eds.), IPMU (2), in: Commun. Comput. Inf. Sci., vol. 443, Springer, 2014, pp. 405–413.
[61] F.E. Boran, D. Akay, A generic method for the evaluation of interval type-2 fuzzy linguistic summaries, IEEE Trans. Cybern. 44 (9) (2014) 1632–1645.
[62] F. Boran, D. Akay, R. Yager, A probabilistic framework for interval type-2 fuzzy linguistic summarization, IEEE Trans. Fuzzy Syst. 22 (6) (2014) 1640–1653.
[63] R.J. Almeida, M.-J. Lesot, B. Bouchon-Meunier, U. Kaymak, G. Moyse, Linguistic summaries of categorical time series for septic shock patient data, in: FUZZ–IEEE, IEEE, 2013, pp. 1–8.
[64] G. Moyse, M.-J. Lesot, B. Bouchon-Meunier, Linguistic summaries for periodicity detection based on mathematical morphology, in: FOCI, IEEE, 2013, pp. 106–113.
[65] G. Moyse, M.-J. Lesot, B. Bouchon-Meunier, Mathematical morphology tools to evaluate periodic linguistic summaries, in: H.L. Larsen, M.J. Martín-Bautista, M.A. Vila, T. Andreasen, H. Christiansen (Eds.), FQAS, in: Lect. Notes Comput. Sci., vol. 8132, Springer, 2013, pp. 257–268.
[66] D. Anderson, R.H. Luke III, J.M. Keller, M. Skubic, M. Rantz, M. Aud, Linguistic summarization of video for fall detection using voxel person and fuzzy logic, Comput. Vis. Image Underst. 113 (1) (2009) 80–89.
[67] D. Anderson, R.H. Luke III, J.M. Keller, M. Skubic, M. Rantz, M. Aud, Modeling human activity from voxel person using fuzzy logic, IEEE Trans. Fuzzy Syst. 17 (1) (2009) 39–49.
[68] M. Ros, M.P. Cuéllar, M. Delgado, A. Vila, D.T. Anderson, J.M. Keller, M. Popescu, Linguistic summarization of long-term trends for understanding change in human behavior, in: FUZZ–IEEE, IEEE, 2011, pp. 2080–2087.
[69] F. Portet, E. Reiter, A. Gatt, J. Hunter, S. Sripada, Y. Freer, C. Sykes, Automatic generation of textual summaries from neonatal intensive care data, Artif. Intell. 173 (7–8) (2009) 789–816.
[70] F. Díaz-Hermida, A. Bugarín, S. Barro, Definition and classification of semi-fuzzy quantifiers for the evaluation of fuzzy quantified sentences, Int. J. Approx. Reason. 34 (1) (2003) 49–88.
[71] V. Novák, I. Perfilieva, M. Holcapek, V. Kreinovich, Filtering out high frequencies in time series using f-transform, Inf. Sci. 274 (2014) 192–209.
[72] V. Novák, I. Perfilieva, On the semantics of perception-based fuzzy logic deduction, Int. J. Intell. Syst. 19 (11) (2004) 1007–1031.
[73] J. Kacprzyk, A. Wilbik, S. Zadrozny, Linguistic summaries of time series via a quantifier based aggregation using the Sugeno integral, in: FUZZ–IEEE, IEEE, 2006, pp. 713–719.
[74] J. Kacprzyk, A. Wilbik, S. Zadrozny, A linguistic approach to a human-consistent summarization of time series using a som learned with a lvq-type algorithm, in: S.D. Kollias, A. Stafylopatis, W. Duch, E. Oja (Eds.), ICANN (2), in: Lect. Notes Comput. Sci., vol. 4132, Springer, 2006, pp. 171–180.
[75] A. Wilbik, J.M. Keller, A fuzzy measure similarity between sets of linguistic summaries, IEEE Trans. Fuzzy Syst. 21 (1) (2013) 183–189.
[76] G. Triviño, A. van der Heide, An experiment on the description of sequences of fuzzy perceptions, in: F. Xhafa, F. Herrera, A. Abraham, M. Köppen, J.M. Benítez (Eds.), HIS, IEEE Computer Society, 2008, pp. 228–233.
[77] R. Castillo-Ortega, N. Marín, C. Martínez-Cruz, D. Sánchez, A proposal for the hierarchical segmentation of time series application to trend-based linguistic description, in: Fuzz–IEEE 2014, 2014.
[78] R. Castillo-Ortega, N. Marín, D. Sánchez, A. Tettamanzi, A multi-objective memetic algorithm for the linguistic summarization of time series, in: GECCO 2011, 2011, pp. 171–172.

[79] R. Castillo-Ortega, N. Marín, D. Sánchez, Linguistic summarization of time series data using genetic algorithms, in: EUSFLAT 2011, 2011, pp. 416–423.
[80] J. Cadenas, N. Marín, M. Vila, Context-aware fuzzy databases, Appl. Soft Comput. 25 (2014) 215–233.
[81] E. Reiter, Nlg vs. templates, in: Proceedings of the Fifth European Workshop on Natural Language Generation, 1995, pp. 95–106.
[82] K. van Deemter, M. Theune, E. Krahmer, Real versus template-based natural language generation: a false opposition?, Comput. Linguist. 31 (1) (2005) 15–24.
[83] L. Eciolaza, M. Pereira-Fariña, G. Triviño, Automatic linguistic reporting in driving simulation environments, Appl. Soft Comput. 13 (9) (2013) 3956–3967.
[84] D. Sanchez-Valdes, L. Eciolaza, G. Triviño, Linguistic description of human activity based on mobile phone's accelerometers, in: J. Bravo, R. Hervás, M. Rodríguez (Eds.), IWAAL, in: Lect. Notes Comput. Sci., vol. 7657, Springer, 2012, pp. 346–353.
[85] D. Sanchez-Valdes, G. Triviño, Computational perceptions of uninterpretable data. A case study on the linguistic modeling of human gait as a quasi-periodic phenomenon, Fuzzy Sets Syst. 253 (2014) 101–121.
[86] A. Alvarez-Alvarez, G. Triviño, Linguistic description of the human gait quality, Eng. Appl. Artif. Intell. 26 (1) (2013) 13–23.
[87] A. Alvarez-Alvarez, D. Sanchez-Valdes, G. Triviño, Á. Sánchez, P.D. Suárez, Automatic linguistic report of traffic evolution in roads, Expert Syst. Appl. 39 (12) (2012) 11293–11302.
[88] D. Anderson, R.H. Luke III, J.M. Keller, M. Skubic, Extension of a soft-computing framework for activity analysis from linguistic summarizations of video, in: FUZZ–IEEE, IEEE, 2008, pp. 1404–1410.
[89] M. Delgado, M. Ros, M.A.V. Miranda, Correct behavior identification system in a tagged world, Expert Syst. Appl. 36 (6) (2009) 9899–9906.
[90] L. González-Villanueva, A. Alvarez-Alvarez, L. Ascari, G. Triviño, A tool for linguistic assessment of rehabilitation exercises, Appl. Soft Comput. 14 (2014) 120–131.
[91] J. Kacprzyk, A. Wilbik, A comprehensive comparison of time series described by linguistic summaries and its application to the comparison of performance of a mutual fund and its benchmark, in: FUZZ–IEEE, IEEE, 2010, pp. 1–8.
[92] D. Sanchez-Valdes, G. Triviño, Increasing the granularity degree in linguistic descriptions of quasi-periodic phenomena, in: H.L. Larsen, M.J. Martín-Bautista, M.A. Vila, T. Andreasen, H. Christiansen (Eds.), FQAS, in: Lect. Notes Comput. Sci., vol. 8132, Springer, 2013, pp. 281–292.
[93] G. Triviño, A. Sanchez, A.S. Montemayor, J.J. Pantrigo, R. Cabido, E.G. Pardo, Linguistic description of traffic in a roundabout, in: FUZZ–IEEE, IEEE, 2010, pp. 1–8.
[94] A. van der Heide, G. Triviño, Automatically generated linguistic summaries of energy consumption data, in: ISDA, IEEE Computer Society, 2009, pp. 553–559.
[95] R. Castillo-Ortega, N. Marín, D. Sánchez, Fuzzy quantification-based linguistic summaries in data cubes with hierarchical fuzzy partition of time dimension, in: E. Corchado, H. Yin (Eds.), 10th International Conference on Intelligent Data Engineering and Automated Learning – IDEAL 2009, Proceedings, Burgos, Spain, September 23–26, 2009, in: Lect. Notes Comput. Sci., vol. 5788, Springer, 2009, pp. 578–585.
[96] A. Ramos-Soto, A. Bugarin, S. Barro, J. Taboada, Linguistic descriptions for automatic generation of textual short-term weather forecasts on real prediction data, IEEE Trans. Fuzzy Syst. 23 (1) (2015) 44–57.
[97] P. Grice, Logic and conversation, in: P. Cole, J. Morgan (Eds.), Syntax and Semantics. 3: Speech Acts, Academic Press, New York, 1975, pp. 41–58.
[98] J. Kacprzyk, A. Wilbik, S. Zadrozny, Mining time series data via linguistic summaries of trends by using a modified Sugeno integral based aggregation, in: CIDM, IEEE, 2007, pp. 742–749.
[99] J. Kacprzyk, A. Wilbik, S. Zadrozny, Linguistic summarization of time series by using the Choquet integral, in: P. Melin, O. Castillo, L.T. Aguilar, J. Kacprzyk, W. Pedrycz (Eds.), IFSA (1), in: Lect. Notes Comput. Sci., vol. 4529, Springer, 2007, pp. 284–294.
[100] J. Kacprzyk, A. Wilbik, S. Zadrozny, Linguistic summaries of time series via an Owa operator based aggregation of partial trends, in: FUZZ–IEEE, IEEE, 2007, pp. 1–6.
[101] F. Díaz-Hermida, A. Ramos-Soto, A. Bugarín, On the role of fuzzy quantified statements in linguistic summarization of data, in: S. Ventura, A. Abraham, K.J. Cios, C. Romero, F. Marcelloni, J.M. Benítez, E.L.G. Galindo (Eds.), ISDA, IEEE, 2011, pp. 166–171.
[102] M. Delgado, M.D. Ruiz, D. Sánchez, M.A. Vila, Fuzzy quantification: a state of the art, Fuzzy Sets Syst. 242 (2014) 1–30.
[103] J. Kacprzyk, A. Wilbik, Linguistic summarization of time series using linguistic quantifiers: augmenting the analysis by a degree of fuzziness, in: FUZZ–IEEE, IEEE, 2008, pp. 1146–1153.
[104] A. Wilbik, J. Kacprzyk, On the evaluation of the linguistic summarization of temporally focused time series using a measure of informativeness, in: IMCSIT, 2010, pp. 155–162.
[105] A. Gatt, F. Portet, Text content and task performance in the evaluation of a natural language generation system, in: G. Angelova, K. Bontcheva, R. Mitkov, N. Nicolov, N. Nikolov (Eds.), RANLP, RANLP 2009 Organising Committee/ACL, 2009, pp. 107–112.
[106] J. Kacprzyk, A. Wilbik, Linguistic summaries of time series using a degree of appropriateness as a measure of interestingness, in: ISDA, IEEE Computer Society, 2009, pp. 385–390.
[107] J. Kacprzyk, A. Wilbik, A new insight into the linguistic summarization of time series via a degree of support: elimination of infrequent patterns, in: D. Dubois, M.A. Lubiano, H. Prade, M.A. Gil, P. Grzegorzewski, O. Hryniewicz (Eds.), SMPS, in: Adv. Soft Comput., vol. 48, Springer, 2008, pp. 393–400.
[108] S. Sripada, E. Reiter, L. Hawizy, Evaluating an nlg system using post-editing, in: L.P. Kaelbling, A. Saffiotti (Eds.), IJCAI, Professional Book Center, 2005, pp. 1700–1701.
[109] J. Kacprzyk, A. Wilbik, Using fuzzy linguistic summaries for the comparison of time series: an application to the analysis of investment fund quotations, in: J.P. Carvalho, D. Dubois, U. Kaymak, J.M. da Costa Sousa (Eds.), IFSA/EUSFLAT Conf., 2009, pp. 1321–1326.
[110] F.P. Romero, J.M. García, Landscapes description using linguistic summaries and a two-dimensional cellular automaton, in: H.L. Larsen, M.J. Martín-Bautista, M.A. Vila, T. Andreasen, H. Christiansen (Eds.), FQAS, in: Lect. Notes Comput. Sci., vol. 8132, Springer, 2013, pp. 221–232.

[111] J. Yu, E. Reiter, J. Hunter, C. Mellish, Choosing the content of textual summaries of large time-series data sets, Nat. Lang. Eng. 13 (1) (2007) 25–49.

[112] R. Castillo-Ortega, N. Marín, D. Sánchez, C. Molina, Flexible querying with linguistic f-cube factory, in: H.L. Larsen, M.J. Martín-Bautista, M.A. Vila, T. Andreasen, H. Christiansen (Eds.), Flexible Query Answering Systems – 10th International Conference, Proceedings, FQAS 2013, Granada, Spain, September 18–20, 2013, in: Lect. Notes Comput. Sci., vol. 8132, Springer, 2013, pp. 245–256.

[113] L. Geng, H. Hamilton, Interestingness measures for data mining: a survey, ACM Comput. Surv. 38 (3) (2006) 1–32, article 9.

[114] E. Hüllermeier, S. Vanderlooy, Why fuzzy decision trees are good rankers, IEEE Trans. Fuzzy Syst. 17 (6) (2009) 1233–1244.

[115] E. Hüllermeier, J. Fürnkranz, Editorial: preference learning and ranking, Mach. Learn. 93 (2–3) (2013) 185–189.

[116] G. Szarvas, R. Busa-Fekete, E. Hüllermeier, Learning to rank lexical substitutions, in: Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing, EMNLP 2013, 18–21 October 2013, Grand Hyatt Seattle, Seattle, Washington, USA, ACL, 2013, pp. 1926–1932, a Meeting of SIGDAT, a Special Interest Group of the ACL.

[117] C. Martínez-Cruz, A. van der Heide, D. Sánchez, G. Triviño, An approximation to the computational theory of perceptions using ontologies, Expert Syst. Appl. 39 (10) (2012) 9494–9503.

[118] M.C. Garrido, N. Marín, O. Pons, Fuzzy intervals to represent fuzzy valid time in a temporal relational database, Int. J. Uncertain. Fuzziness Knowl.-Based Syst. 17 (Supplement-1) (2009) 173–192.