Linguistic Summarization of Time Series Under Different Granulation of Describing Features

Janusz Kacprzyk, Anna Wilbik, and Sławomir Zadrożny

Systems Research Institute, Polish Academy of Sciences
ul. Newelska 6, 01-447 Warsaw, Poland
{kacprzyk, wilbik, zadrozny}@ibspan.waw.pl

Abstract. We consider an extension to a new approach to the linguistic summarization of time series data proposed in our previous papers. We summarize trends identified here with straight segments of a piecewise linear approximation of time series. Then we employ, as a set of features, the duration, dynamics of change and variability, and assume different, human consistent granulations of their values. The problem boils down to a linguistic quantifier driven aggregation of partial trends that is done via the classic Zadeh's calculus of linguistically quantified propositions but with different t-norms. We show an application to linguistic summarization of time series data on daily quotations of an investment fund over an eight year period.

1 Introduction

A linguistic data (base) summary, meant as a concise, human-consistent description of a (numerical) data set, was introduced by Yager [18] and then further developed by Kacprzyk and Yager [11], and Kacprzyk, Yager and Zadrożny [12]. The contents of a database is summarized via a natural language like expression semantics provided in the framework of Zadeh's calculus of linguistically quantified propositions [21]. Since data sets are usually large, it is very difficult for a human being to capture and understand their contents. As natural language is the only fully natural means of articulation and communication for a human being, such linguistic descriptions are the most human consistent.

In this paper we consider a specific type of data, namely time series. In this context it might be good to obtain a brief, natural language like description of trends present in the data on, e.g., stock exchange quotations, sales, etc. over a certain period of time.

Though statistical methods are widely used, we wish to derive (quasi)natural language descriptions to be considered to be an additional form of data description of a remarkably high human consistency. Hence, our approach is not meant to replace the classical statistical analyses but to add a new quality.

The summaries of time series we propose refer in fact to the summaries of trends identified here with straight line segments of a piece-wise linear approximation of time series. Thus, the first step is the construction of such an approximation. For this purpose we use a modified version of the simple, easy to use Sklansky and Gonzalez algorithm presented in [16].

M. Kryszkiewicz et al. (Eds.): RSEISP 2007, LNAI 4585, pp. 230–240, 2007.

Then we employ a set of features (attributes) to characterize the trends such as the slope of the line, the fairness of approximation of the original data points by line segments and the length of a period of time comprising the trend.

Basically the summaries proposed by Yager are interpreted in terms of the number or proportion of elements possessing a certain property. In the framework considered here a summary might look like: “Most of the trends are short” or in a more sophisticated form: “Most long trends are increasing”. Such expressions are easily interpreted using Zadeh’s calculus of linguistically quantified propositions. The most important element of this interpretation is a linguistic quantifier exemplified by “most”. In Zadeh’s [21] approach it is interpreted in terms of a proportion of elements possessing a certain property (e.g., a length of a trend) among all the elements considered (e.g., all trends).

In Kacprzyk, Wilbik and Zadrożny [6] we proposed to use Yager’s linguistic summaries, interpreted in the framework of Zadeh’s calculus of linguistically quantified propositions, for the summarization of time series. In our further papers (cf. Kacprzyk, Wilbik and Zadrożny [8,9,10]) we extended this idea by proposing other types of summaries and the use of other mathods, notably the Choquet and Sugeno integrals. All these approaches have been proposed using a unified perspective given by Kacprzyk and Zadrożny [13] that is based on Zadeh’s [22] protoforms.

In this paper we employ the classic Zadeh’s calculus of linguistically quantified propositions. However, we will extend the idea proposed in our source paper (Kacprzyk, Wilbik and Zadrożny [6]) by using various t-norms and show results of an application to data on daily quotations of a mutual (investment) fund over an eight year period.

The paper is in line with some modern approaches to a human consistent summarization of time series – cf. Batyrshin and his collaborators [1,2], or Chiang, Chow and Wang [4] but we use a different approach.

One should mention an interesting project coordinated by the University of Aberdeen, UK, `SumTime`, an EPSRC Funded Project for Generating Summaries of Time Series Data[1]. Its goal is also to develop a technology for producing English summary descriptions of a time-series data set using an integration of time-series and natural language generation technology. Linguistic summaries obtained related to wind direction and speed are, cf. Sripada et al. [17]:

- WSW (West of South West) at 10-15 knots increasing to 17-22 knots early morning, then gradually easing to 9-14 knots by midnight,
- During this period, spikes simultaneously occur around 00:29, 00:54, 01:08, 01:21, and 02:11 (o’clock) in these channels.

They do provide a higher human consistency as natural language is used but they capture imprecision of natural language to a very limited extent. In our approach this will be overcome to a considerable extent.

[1] `www.csd.abdn.ac.uk/research/sumtime/`

2 Temporal Data and Trend Analysis

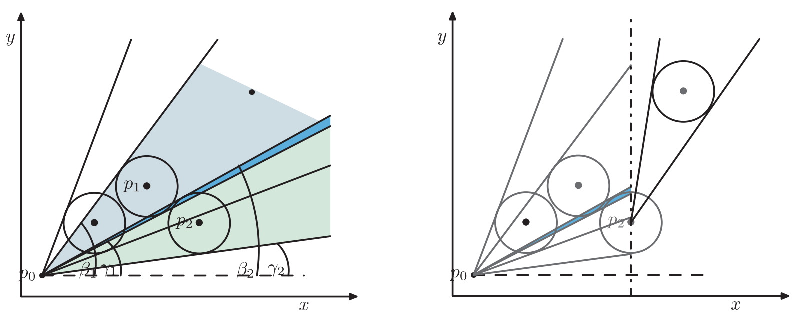
We identify trends as linearly increasing, stable or decreasing functions, and therefore represent given time series data as piecewise linear functions of some slope (intensity of an increase and decrease). These are partial trends as a global trend concerns the entire time span. There also may be trends that concern more than a window taken into account while extracting partial trends by using the Sklansky and Gonzalez [16] algorithm.

We use the concept of a uniform partially linear approximation of a time series. Function f is a uniform ε-approximation of a set of points $\{(x_i, y_i)\}$, if for a given, context dependent $\varepsilon > 0$, there holds

$$\forall i: \ |f(x_i) - y_i| \leq \varepsilon \tag{1}$$

and if f is linear, then such an approximation is a linear uniform ε-approximation.

We use a modification of the well known Sklansky and Gonzalez [16] algorithm that finds a linear uniform ε-approximation for subsets of points of a time series. The algorithm constructs the intersection of cones starting from point p_i of the time series and including a circle of radius ε around the subsequent points p_{i+j}, $j = 1, 2, \ldots,$ until the intersection of all cones starting at p_i is empty. If for p_{i+k} the intersection is empty, then we construct a new cone starting at p_{i+k-1}. Figures 1(a) and 1(b) present the idea of the algorithm. The family of possible solutions is indicated as a gray area. For other algorithms, see,e.g., [15].



(a) the intersection of the cones is indicated by the dark grey area

(b) a new cone starts in point p_2

Fig. 1. An illustration of the algorithm for the uniform ε-approximation

First, denote:`p_0` – a point starting the current cone, `p_1` – the last point checked in the current cone, `p_2` – the next point to be checked, `Alpha_01` – a pair of angles (γ_1, β_1), meant as an interval, that defines the current cone as in Fig. 1(a), `Alpha_02` – a pair of angles of the cone starting at `p_0` and inscribing the circle of radius ε around `p_2` (cf. (γ_2, β_2) in Fig. 1(a)), function `read_point()` reads a next point of data series, function `find()` finds a pair of

```
read_point(p_0);
read_point(p_1);
while(1)
{
  p_2=p_1;
  Alpha_02=find();
  Alpha_01=Alpha_02;
  do
  {

    Alpha_01 = Alpha_01 ∩ Alpha_02;

    p_1=p_2;
    read_point(p_2);
    Alpha_02=find();

  } while(Alpha_01 ∩ Alpha_02 ≠ ∅);

  save_found_trend();
  p_0=p_1;
  p_1=p_2;
}
```

Fig. 2. Pseudocode of the modified Sklansky and Gonzalez [16] algorithm for extracting trends

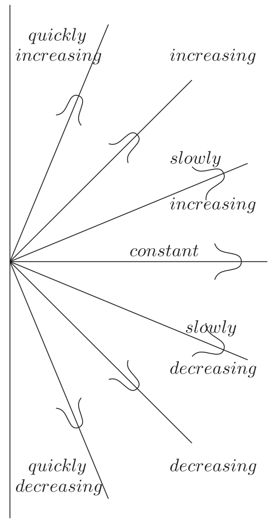


Fig. 3. A visual representation of angle granules defining the dynamics of change

angles of the cone starting at `p_0` and inscribing the circle of radius ε around `p_2`. Then, a pseudocode of the algorithm that extracts trends is given in Fig. 2.

The bounding values of `Alpha_02` (γ_2, β_2), computed by function `find()` correspond to the slopes of two lines tangent to the circle of radius ε around $p_2 = (x_2, y_2)$ and starting at $p_0 = (x_0, y_0)$. Thus, if $\Delta x = x_0 - x_2$ and $\Delta y = y_0 - y_2$ then:

$$\gamma_2 = arctg\left[\left(\Delta x \cdot \Delta y \pm \varepsilon\sqrt{(\Delta x)^2 + (\Delta y)^2 - \varepsilon^2}\right) / \left((\Delta x)^2 - \varepsilon^2\right)\right]$$

The resulting linear ε-approximation of a group of points `p_0, ... ,p_1` is either a single segment, chosen as, e.g., a bisector of the cone, or one that minimizes the distance (e.g., the sum of squared errors, SSE) from the approximated points, or the whole family of possible solutions, i.e., the rays of the cone.

3 Dynamic Characteristics of Trends

While summarizing trends in time series data, we consider the following three aspects: (1) dynamics of change, (2) duration, and (3) variability, and by trends we mean here global trends, concerning the entire time series (or some, probably a large, part of it), not partial trends concerning in the (partial) trend extraction phase via the Sklansky and Gonzales [16] algorithm. In what follows we will briefly discuss these factors.

Dynamics of change

By *dynamics of change* we understand the speed of changes. It can be described by the slope of a line representing the trend, (cf. any angle η from the interval $\langle\gamma, \beta\rangle$ in Fig. 1(a)). Thus, to quantify dynamics of change we may use the interval of possible angles $\eta \in \langle -90; 90\rangle$.

For practical reasons, we use a fuzzy granulation via a scale of linguistic terms as, e.g.: *quickly decreasing*, *decreasing*, *slowly decreasing*, *constant*, *slowly increasing*, *increasing*, *quickly increasing*, as illustrated in Fig. 3. Batyrshin et al. [1,2] give some methods for constructing such a fuzzy granulation.

We map a single value α (or the interval of angles corresponding to the gray area in Fig. 1(b)) characterizing the dynamics of change into a fuzzy set (linguistic label) best matching a given angle, and we can say that a given trend is, e.g., "decreasing to a degree 0.8".

Duration

Duration describes the length of a single trend, meant as a linguistic variable and exemplified by a "long trend" defined as a fuzzy set.

Variability

Variability refers to how "spread out" (in the sense of values) a group of data is. Traditionally, the following five statistical measures of variability are widely used:

- The range (maximum – minimum).
- The interquartile range (IQR) calculated as the third quartile (the 75th percentile) minus the first quartile (the 25th percentile) that may be interpreted as representing the middle 50% of the data.
- The variance is calculated as $\sum_i (x_i - \bar{x})^2/n$, where $\bar{x}$ is the mean value.
- The standard deviation – a square root of the variance.
- The mean absolute deviation (MAD), calculated as $\sum_i |x_i - \bar{x}|/n$.

We measure the variability of a trend as the distance of the data points from its linear uniform ε-approximation (cf. Section 2). We propose to employ a distance between a point and a family of possible solutions, indicated as a gray cone in Fig. 1(a). Equation (1) assures that the distance is definitely smaller than ε. The normalized distance equals 0 if the point lays in the gray area and otherwise is equal to the distance to the nearest point belonging to the cone, divided by ε.

Then, we find for a given value of variability obtained a best matching fuzzy set (linguistic label).

4 Linguistic Data Summaries

A linguistic summary is meant as a (short) natural language like sentence(s) that subsumes the very essence of a (numeric, usually large) set of data (cf. Kacprzyk and Zadrożny [13], [14]). In Yager's approach (cf. Yager [18], Kacprzyk and Yager [11], and Kacprzyk, Yager and Zadrożny [12]) the following perspective for linguistic data summaries is assumed:

– $Y = \{y_1, \ldots, y_n\}$ is a set of objects in a database, e.g., the set of workers;
– $A = \{A_1, \ldots, A_m\}$ is a set of attributes characterizing objects from Y, e.g., salary, and $A_j(y_i)$ is a value of attribute A_j for object y_i.

A linguistic summary of a data set consists of:

– a summarizer P, i.e. an attribute together with a linguistic value (fuzzy predicate) defined on the domain of attribute A_j (e.g. "low" for attribute "salary");
– a quantity in agreement Q, i.e. a linguistic quantifier (e.g. most);
– truth (validity) $\mathcal{T}$ of the summary, i.e. a number from the interval $[0, 1]$ assessing the truth (validity) of the summary (e.g. 0.7); usually, only summaries with a high value of $\mathcal{T}$ are interesting;
– optionally, a qualifier R, i.e. another attribute together with a linguistic value (fuzzy predicate) defined on the domain of attribute A_k determining a (fuzzy subset) of Y (e.g. "young" for attribute "age").

Thus, a linguistic summary may be exemplified by

$$\mathcal{T}(\text{most of employees earn low salary}) = 0.7 \tag{2}$$

or, in a richer (extended) form, including a qualifier (e.g. young), by

$$\mathcal{T}(\text{most of young employees earn low salary}) = 0.9 \tag{3}$$

Thus, basically, the core of a linguistic summary is a *linguistically quantified proposition* in the sense of Zadeh [21] which, for (2) and (3), respectively, may be written as

$$Qy\text{'s are } P \qquad\qquad QRy\text{'s are } P \tag{4}$$

Then, $\mathcal{T}$ directly corresponds to the truth value of (4) that may be calculated by Zadeh's calculus of linguistically quantified propositions (cf. [21] or the next section), or other interpretations of linguistic quantifiers (cf. [7]).

5 Protoforms of Linguistic Trend Summaries

As shown by Kacprzyk and Zadrożny [13], Zadeh's [22] concept of a protoform is convenient for dealing with linguistic summaries. A protoform is defined as a more or less abstract prototype (template) of a linguistically quantified proposition. Then, the summaries mentioned above may be represented by two types of the protoforms:

– a protoform of a short form of linguistic summaries:

$$Q \text{ trends are } P \tag{5}$$

and exemplified by: *Most* of trends are of a *large variability*

– a protoform of an extended form of linguistic summaries:

$$QR \text{ trends are } P \tag{6}$$

and exemplified by: *Most* of *slowly decreasing trends* are of a *large variability*

Their truth values will be found using the classic Zadehs calculus of linguistically quantified propositions as it is effective and efficient, and provides the best conceptual framework for a linguistic quantifier driven aggregation of partial trends.

6 The Use of Zadeh's Calculus

Using Zadeh's [21] fuzzy logic based calculus of linguistically quantified propositions, a (proportional, nondecreasing) linguistic quantifier Q is assumed to be a fuzzy set defined, i.e. $\mu_Q : [0,1] \longrightarrow [0,1]$, $\mu_Q(x) \in [0,1]$. We consider *regular non-decreasing monotone* quantifiers, as e.g. "most" given by (8):

$$\mu(0) = 0; \qquad \mu(1) = 1; \qquad x_1 \leq x_2 \Rightarrow \mu_Q(x_1) \leq \mu_Q(x_2) \tag{7}$$

$$\mu_Q(x) = \begin{cases} 1 & \text{for } x \underline{>} 0.8 \\ 2x - 0.6 & \text{for } 0.3 < x < 0.8 \\ 0 & \text{for } x \underline{<} 0.3 \end{cases} \tag{8}$$

The truth values (from [0,1]) of (5) and (6) are calculated, respectively, as

$$\mathcal{T}(Qy\text{'s are } P) = \mu_Q\left(\frac{1}{n}\sum_{i=1}^{n} \mu_P(y_i)\right) \tag{9}$$

$$\mathcal{T}(QRy\text{'s are } P) = \mu_Q\left(\frac{\sum_{i=1}^{n}(\mu_R(y_i) \wedge \mu_P(y_i))}{\sum_{i=1}^{n}\mu_R(y_i)}\right) \tag{10}$$

where $\wedge$ is the minimum (more generally, e.g., a t-norm).

Both the fuzzy predicates P and R are assumed of a simplified, atomic form referring to one attribute, but can be extended to cover some confluences of various, multiple attribute values.

A *t-norm* is a $t : [0,1] \times [0,1] \longrightarrow [0,1]$, such that, for each $a, b, c \in [0,1]$:

1. it has 1 as the unit element, i.e. $t(a,1) = a$,
2. it is monotone, i.e. $a \leq b \Longrightarrow t(a,c) \leq t(b,c)$,
3. it is commutative, i.e. $t(a,b) = t(b,a)$,
4. it is associative, i.e. $t[a, t(b,c)] = t[t(a,b), c]$.

Some more relevant examples of t-norms are: (1) the minimum $t(a,b) = a \wedge b = \min(a,b)$ which is the most widely used, also here, (2) the algebraic product $t(a,b) = a \cdot b$, (3) the Łukasiewicz t-norm $t(a,b) = \max(0, a+b-1)$, and (4) the drastic t-norm $t(a,b) = \begin{cases} b \ \ a = 1 \\ a \ \ b = 1 \\ 0 \ \ otherwise \end{cases}$.

These operations can be in principle used in Zadeh's calculus but, clearly, their use may result in different results of the linguistic quantifier driven aggregation. Some examples will be shown in the next section.

7 Numerical Experiments

The method was tested on real data of daily quotations, from April 1998 to December 2006, of an investment fund that invests at most 50% of assets in shares, cf. Fig. 4, with the starting value of one share equal to PLN 10.00 and the final one equal to PLN 45.10 (PLN stands for the Polish Zloty); the minimum was PLN 6.88 while the maximum was PLN 45.15, and the biggest daily increase was PLN 0.91, while the biggest daily decrease was PLN 2.41.

For $\varepsilon = 0.25$ (PLN 0.25), we obtained 255 extracted trends, ranging from 2 to 71 time units (days). The histogram of duration is in Fig. 5.

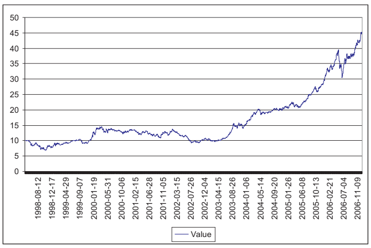


Fig. 4. A view of the original data

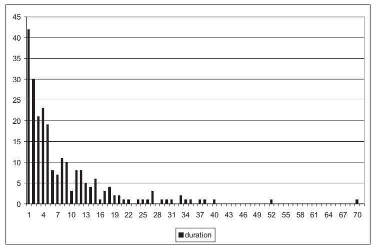


Fig. 5. Histogram of duration of trends

Figure 6 shows the histogram of angles (dynamics of change) and the histogram of variability of trends (in %) is in Fig. 7.

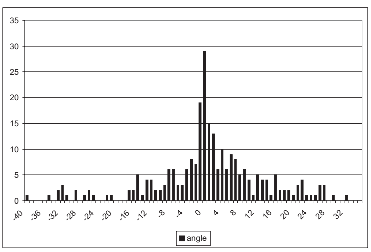


Fig. 6. Histogram of angles decscribing dynamic of change

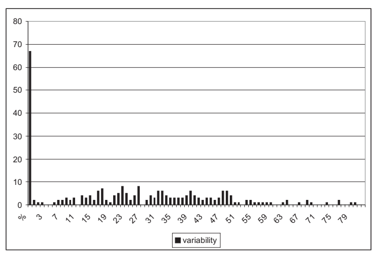


Fig. 7. Histogram of variability of trends

Some interesting summaries obtained, for different granulations of the dynamics of change, duration and variability, are:

- for 7 labels for the dynamics of change (*quickly increasing, increasing, slowly increasing, constant, slowly decreasing, decreasing and quickly decreasing*), 5 labels for the duration (*very long, long, medium, short, very short*) and 5 labels the variability (*very high, high, medium, low, very low*):

  - Most trends are very short, $\mathcal{T} = 0.78$
  - for different t-norms are shown in Table 1.
- 5 labels for the dynamics of change (*increasing, slowly increasing, constant, slowly decreasing, decreasing*), 3 labels for the duration (*short, medium, long*) and 5 labels for the variability (*very high, high, medium, low, very low*):
  - Most trends are of medium length, $\mathcal{T} = 0.431$
  - for different t-norms are shown in Table 2.

Table 1. Truth values for extended form summaries with different t-norms for the first granulation

Summary	minimum	product	Łukasiewicz	drastic
Most trends with a low variability are constant	0.974	0.944	0.911	0.85
Most slowly decreasing trends are of a very low variability	0.636	0.631	0.63	0.589
Almost all short trends are constant	1	1	1	1

Table 2. Truth values for extended form summaries with different t-norms for the second granulation

Summary	minimum	product	Łukasiewicz	drastic
Almost all decreasing trends are short	1	1	1	1
Almost all increasing trends are short	0.58	0.514	0.448	0.448
At least a half of medium length trends are constant	0.891	0.877	0.863	0.863
Most of slowly increasing trends are of a medium length	0.798	0.773	0.748	0.748
Most of trends with a low variability are constant	0.567	0.517	0.466	0.466
Most of trends with a very low variability are short	0.909	0.9	0.891	0.891
Most trends with a high variability are of a medium length	0.801	0.754	0.707	0.707
None of trends with a very high variability is long	1	1	1	1
None of decreasing trends is long	1	1	1	1
None of increasing trends is long	1	1	1	1

The particular linguistic summaries obtained, and their associated truth values, are intuitively appealing. In addition, these summaries were found interesting by domain experts though a detailed analysis from the point of view of financial analyses is beyond the scope of this paper. The results obtained for different t-norms are similar and, of course, the truth value for the case of the minimum is the highest.

8 Concluding Remarks

We proposed new types of lingustic summaries of time series. The derivation of a linguistic summary of a time series was related to a liguistic quantifier driven aggregation of trends, and we employed the classic Zadeh's calculus of linguistically quantified propositions with different t-norms, not only the classic minimum. We showed an application to the analysis of time series data on daily quotations of an investment fund over an eight year period, present some interesting lingustic sumaries obtained, and showed results for different t-norms. They suggest that varous t-norms exhibit slightly different behavior and they choice may be crucial for a particular application. The results are very promising.

References

1. Batyrshin, I.: On granular derivatives and the solution of a granular initial value problem. International Journal Applied Mathematics and Computer Science 12(3), 403–410 (2002)
2. Batyrshin, I., Sheremetov, L.: Perception based functions in qualitative forecasting. In: Batyrshin, I., Kacprzyk, J., Sheremetov, L., Zadeh, L.A. (eds.) Perception-based Data Mining and Decision Making in Economics and Finance, Springer, Heidelberg (2006)
3. Berndt, D.J., Clifford, J.: Finding patterns in time series: a dynamic programming approach. In: Advances in Knowledge Discovery and Data Mining, pp. 229–248. AAAI/MIT Press, Menlo Park, CA (1996)
4. Chiang, D.-A., Chow, L.R., Wang, Y.-F.: Mining time series data by a fuzzy linguistic summary system. Fuzzy Sets and Systems 112, 419–432 (2000)
5. Das, G., Lin, K., Mannila, H., Renganathan, G., Smyth, P.: Rule discovery from time series. In: Proc. of the 4th Int'l Conference on Knowledge Discovery and Data Mining. New York, NY, pp. 16–22 (1998)
6. Kacprzyk, J., Wilbik, A., Zadrożny, S.: Linguistic summarization of trends: a fuzzy logic based approach. In: Proceedings of the 11th International Conference Information Processing and Management of Uncertainty in Knowledge-based Systems, Paris, France, July 2-7, 2006, pp. 2166–2172 (2006)
7. Kacprzyk, J., Wilbik, A., Zadrożny, S.: Linguistic summaries of time series via a quantifier based aggregation using the Sugeno integral. In: Kacprzyk, J., Wilbik, A. (eds.) Proceedings of 2006 IEEE World Congress on Computational Intelligence, Vancouver, BC, Canada, July 16-21, 2006, pp. 3610–3616. IEEE Press, New York (2006)
8. Kacprzyk, J., Wilbik, A., Zadrożny, S.: On some types of linguistic summaries of time series. In: Proceedings of the 3rd International IEEE Conference Intelligent Systems, London, UK, pp. 373–378. IEEE Press, New York (2006)
9. Kacprzyk, J., Wilbik, A., Zadrożny, S.: A linguistic quantifier based aggregation for a human consistent summarization of time series. In: Lawry, J., Miranda, E., Bugarin, A., Li, S., Gil, M.A., Grzegorzewski, P., Hryniewicz, O. (eds.) Soft Methods for Integrated Uncertainty Modelling, pp. 186–190. Springer, Heidelberg (2006)
10. Kacprzyk, J., Wilbik, A., Zadrożny, S.: Capturing the essence of a dynamic behavior of sequences of numerical data using elements of a quasi-natural language. In: Proceedings of the 2006 IEEE International Conference on Systems, Man, and Cybernetics, Taipei, Taiwan, 2006, pp. 3365–3370. IEEE Press, New York (2006)
11. Kacprzyk, J., Yager, R.R.: Linguistic summaries of data using fuzzy logic. International Journal of General Systems 30, 33–154 (2001)
12. Kacprzyk, J., Yager, R.R., Zadrożny, S.: A fuzzy logic based approach to linguistic summaries of databases. International Journal of Applied Mathematics and Computer Science 10, 813–834 (2000)
13. Kacprzyk, J., Zadrożny, S.: Linguistic database summaries and their protoforms: toward natural language based knowledge discovery tools. Information Sciences 173, 281–304 (2005)
14. Kacprzyk, J., Zadrożny, S.: Fuzzy linguistic data summaries as a human consistent, user adaptable solution to data mining. In: Gabrys, B., Leiviska, K., Strackeljan, J. (eds.) Do Smart Adaptive Systems Exist? pp. 321–339. Springer, Heidelberg (2005)

15. Keogh, E., Pazzani, M.: An enhanced representation of time series which allows fast and accurate classification, clustering and relevance feedback. In: Proc. of the 4th Int'l Conference on Knowledge Discovery and Data Mining. New York, NY, pp. 239–241 (1998)
16. Sklansky, J., Gonzalez, V.: Fast polygonal approximation of digitized curves. Pattern Recognition 12(5), 327–331 (1980)
17. Sripada, S., Reiter, E., Davy, I.: SumTime-Mousam: Configurable Marine Weather Forecast Generator. Expert Update 6(3), 4–10 (2003)
18. Yager, R.R.: A new approach to the summarization of data. Information Sciences 28, 69–86 (1982)
19. Yager, R.R.: On ordered weighted averaging aggregation operators in multicriteria decision making. IEEE Transactions on Systems, Man. and Cybernetics SMC 18, 183–190 (1988)
20. Yager, R.R., Kacprzyk, J.: The Ordered Weighted Averaging Operators: Theory and Applications. Kluwer, Boston (1997)
21. Zadeh, L.A.: A computational approach to fuzzy quantifiers in natural languages. Computers and Mathematics with Applications 9, 149–184 (1983)
22. Zadeh, L.A.: A prototype-centered approach to adding deduction capabilities to search engines – the concept of a protoform. In: Proceedings of the Annual Meeting of the North American Fuzzy Information Processing Society (NAFIPS, 2002) pp. 523–525 (2002)
23. Zadeh, L.A., Kacprzyk, J. (eds.): Computing with Words in Information/Intelligent Systems: 1. Foundations, 2. Applications. Physica-Verlag, Heidelberg (1999a)