On a Paradox of Extended Linguistic Summaries

Anna Wilbik
Information Systems
School of Industrial Engineering
Eindhoven University of Technology
Eindhoven, The Netherlands
a.m.wilbik@tue.nl

Timothy C. Havens
College of Computing
Michigan Technological University
Houghton, MI, USA
thavens@mtu.edu

Tim Wilkin
School of Information Technology
Deakin University
Geelong, Australia
tim.wilkin@deakin.edu.au

***Abstract*—Continued developments in information technologies allows for increasingly more data to be collected for decision making purposes. While statistical summaries and aggregations are commonly applied to such data, linguistic summaries capture essential features and relationships in the data and better support human users to understand complex data sets. The basic quality measure of linguistic summaries is the truth value, describing the validity of the sentence. Several methods for calculating the truth value have been proposed. In this paper we analyze several popular methods and show a strange, contradictory behavior in case of extended protoforms, which can result in misleading or non-intuitive results to the user. These results highlight the need for further research into linguistic summarization and the computation of truth values for real data sets.**

***Index Terms*—linguistic summarization, extended protoform, truth value, dual summary, paradox**

I. Introduction

Data drives a vast majority of modern business, industrial, and governance processes. Over the past decades the data has enabled us to gain deep and significant insights into social, political, and economic structures around the world. However, estimates of the continual growth of data volumes generated annually is as high as 61% CAGR [1] and given that humanity is generating approximately 2.5 exabytes of data daily, we produce far more data than can be readily accessed and processed to support important decision processes. While automation of such processes is one important research direction, human decision makers are still required in many critical domains. Decision support systems aide decision makers by summarizing and aggregating data, and transforming it into formats and feature spaces that humans can more readily interpret and understand. While many statistical and machine-driven data analysis methods utilize quantitative data, humans are more naturally disposed to interpret visual and natural language representations of information. While graphs are a common way to visually represent data relationships and insights, they still require training and experience to be able to interpret quickly and accurately [2]. Natural language remains the most salient means of communicating with humans. Computational models for linguistic summarization of data aim to provide human-centered methods for communicating insights and critical information from quantitative data sources.

For instance, Dubois and Prade [3], [4] proposed representation and reasoning for gradual inference rules for linguistic summarization in the form "the more X is F, the more/the less Y is G" that could summarize various relationships. Such rules expressed a progressive change of the degree to which the entity Y satisfies the gradual property G when the degree to which the entity X satisfies the gradual property F is modified. This concept was later investigated by Oudini et. al [5], [6] and Wilbik et. al [7]. Rasmussen and Yager [8] discussed the benefit of using fuzzy sets in data summaries based on generalized association rules. Gradual functional dependencies were investigated, and a query language called SummarySQL was proposed. Bosc et al. [9] discussed the use of fuzzy cardinalities for linguistic summarization. The SAINTETIQ model [10] provides the user synthetic views of groups of tuples over the database. In [11], the authors proposed a summarization procedure to describe long-term trends of change in human behavior, e.g., "the quality of the 'wake up' behavior has been decreasing in the last month" or "the quality of the 'morning routine' is constant but has been highly unstable in the last month."

We will follow the approach of Yager [12], in the form "Q objects in Y are P." This approach was considerably advanced and then implemented by Kacprzyk [13], Kacprzyk and Yager [14], and Kacprzyk et al. [15]–[17]. This approach has also been applied to different types of data: numerical [18]–[20], time series [21]–[23], sensor data [11], [24], texts [25], videos [26]–[28] and processes [29], [30].

In this paper we focus on the linguistic summaries of numerical data and the extended protoform [12], i.e., "Q R objects in Y are P." In our work we have encountered a strange behaviour of linguistic summaries that we describe and investigate herein.

This paper is structured as follows. The next section introduces the background of linguistic summaries and Section III demonstrates the encountered paradox. We then demonstrate how different methods for calculating the truth value behave in this paradoxical situation. Last, are some concluding remarks and recommendations for further study.

This work was supported in part by the Michigan Tech Visiting Professor program, which is funded by a grant to the Michigan Tech Provost's Office from the State of Michigan's King-Chavez-Parks Initiative, and the Michigan Tech Institute of Computing and Cybersystems.

978-1-7281-6932-3/20/$31.00 ©2020 IEEE

II. BACKGROUND

We consider linguistic data summaries as quantified propositions with two possible protoforms (or templates) [12]:

- simple protoform:

$$Qy\text{'s are } P; \quad (1)$$

  e.g. *Most* boxes are *large*
- extended protoform:

$$Q\ Ry\text{'s are } P; \quad (2)$$

  e.g. *Most large* boxes are *heavy*

where Q is the quantifier, P is the summarizer, and R is an optional qualifier, which are all modeled as fuzzy sets over appropriate domains.

The *truth value*, describing the validity of the summary, is the basic measure of the quality of the summary; thus, it is an essential part of the summary. Many methods for calculating the truth value have been proposed; for some examples, see [31]. The truth value is not the only quality measure of a linguistic summary. Kacprzyk et al. [17], [32] proposed four additional measures, namely the degree of specificity, the degree of appropriateness, the degree of covering, and the length of the summary. Bugarin et. al [33] differentiated between evaluating a single summary sentence and a set of summaries. They proposed several measures that capture aspects, such as coverage, length, and specificity. An overview of different quality criteria can be found in [34]. For completeness purposes, we now briefly describe the four methods that we use in this work.

A. Zadeh's calculus of quantified propositions

In this approach [35], the truth value, denoted as $\mathcal{T}_{\mathcal{Z}}$, of a simple protoform summary is calculated as

$$\mathcal{T}_{\mathcal{Z}}(Qy\text{'s are } P) = \mu_Q\left(\frac{1}{n}\sum_{i=1}^{n}\mu_P(y_i)\right); \quad (3)$$

the extended protoform summary is calculated as

$$\mathcal{T}_{\mathcal{Z}}(Q\ Ry\text{'s are } P) = \mu_Q\left(\frac{\sum_{i=1}^{n}\mu_P(y_i)\wedge\mu_R(y_i)}{\sum_{i=1}^{n}\mu_R(y_i)}\right), \quad (4)$$

where n is the number of objects in the data, and μ_P, μ_Q, and μ_R are the membership functions of the summarizer, the qualifier, and the quantifier, respectively. The notation $\wedge$ denotes the minimum operator; in the more general case, it can be any t-norm [36].

B. Method based on Sugeno integral

Using the approach in [20], the truth value, denoted as $\mathcal{T}_{\mathcal{S}}$, is computed by

$$\mathcal{T}_{\mathcal{S}}(Q\ Ry\text{'s are } P) = \mathcal{T}_{\mathcal{S}}(Q_1\ Ry\text{'s are } P)\wedge\mathcal{T}_{\mathcal{S}}(Q_2\ Ry\text{'s are } P) \quad (5)$$

where quantifier Q is split into Q_1 and Q_2, so that Q_1 is a non-increasing quantifier and Q_2 is non-decreasing. For the non-decreasing quantifier Q_2 the truth value is calculated as

$$\mathcal{T}_{\mathcal{S}}(Q_2\ Ry\text{'s are } P) = \max_{\beta\in[0,\max(R(y_i))]}\beta\wedge\left(\max_{\alpha\in[0,1]}\left\{\alpha\wedge Q_2(P_{\alpha^{R_\beta}})\right\}\right), \quad (6)$$

where $R_\beta\{y_i\in Y:\mu_R(y_i)\geq\beta\}$ and

$$P_\alpha^{R_\beta} = \frac{|\{y_i\in R_\beta:\mu_P(y_i)\geq\alpha\}|}{|R_\beta|} \text{ for } |R_\beta|>0.$$

For the non-increasing quantifier Q_1, the truth value is calculated for the dual summary to "$Q_2\ R\ y$'s are P." Namely, "$\widehat{Q_2}\ R\ y$'s are $\overline{P}$," where $\widehat{\cdot}$ denotes antonym and $\overline{\cdot}$ is complement.

C. GD method

The truth value according to the GD method [37] is calculated as

$$GD(Q\ Ry\text{'s are } P) = \sum_{c\in CR(P/R)} ER(P/R,c)\times\mu_Q(c), \quad (7a)$$

where

$$ER(P/R,c) = \sum_{c=C(P/R,\alpha_i)}(\alpha_i-\alpha_{i+1}),\ \forall c\in CR(P/R), \quad (7b)$$

$$CR(P/R) = \left\{\frac{|(P\cap R)_\alpha|}{|R_\alpha|}:\alpha\in M(P/R)\right\}, \quad (7c)$$

$$M(P/R) = M(P\cap R)\cup M(R), \quad (7d)$$

$$M(P) = \{\alpha\in]0,1]:\exists y_i,\mu_P(y_i)=\alpha\}, \quad (7e)$$

$$C(P/R,\alpha_i) = \frac{|(P\cap R)_{\alpha_i}|}{|R_{\alpha_i}|}. \quad (7f)$$

If R is not a normal fuzzy set, it is first normalized. The fuzzy set $P\cap R$ is scaled using the same factor used in the normalization of R.

D. ZS method

The truth value according to the ZS method [37] is calculated as

$$ZS(Q\ Ry\text{'s are } P) = \max_{c\in CR(P,R)}\min\{ES(P,R,c),\mu_Q(c)\} \quad (8)$$

where

$$ES(P/R,c) = \max\left\{\alpha\in M(P/R):c=\frac{|(P\cap R)_\alpha|}{|R_\alpha|}\right\},\ \forall c\in CR(P/R),$$

and $M(P/R)$ and $CR(P/R)$ are defined as in GD method at (7d) and (7c), respectively. Similar to GD, if fuzzy set R is normalized, then the fuzzy set $P\cap R$ is scaled using the same factor used in the normalization of R.

E. Properties for evaluating linguistic summaries

Delgado, Sanchez and Vila [37] proposed a set of properties for evaluating linguistic summaries: six for simple protoforms and eight for extended protoforms. Here, we reintroduce only the properties for extended protoforms.

- (Crisp case). If A and D are crisp, then the (known) result of the evaluation must be $Q\left(\frac{|P \cap R|}{|R|}\right)$, where Q is relative quantifier.
- In the case $R = X$ and for relative quantifiers, the resulting evaluation method is a valid method for the evaluation of simple protoform sentences.
- Evaluation must be time-efficient.
- If $R \subseteq P$ and R is a normal set then the evaluation method must return the value $Q(1)$.
- If $D \cap A = \varnothing$ then the evaluation method must return the value $Q(0)$.
- Evaluation must be coherent with fuzzy logic in the case of the quantifiers "exist" and "all".
- Evaluation must allow us to use any quantifier, i.e. any possibility distribution over $[0, 1]$.
- Evaluation must not be too "strict", i.e. given a quantifier Q in the rational interval [0,1] with $Q \neq \varnothing$ and $Q \neq H = \{p/q \text{ with } p \in \{0, \ldots, n\} \text{ and } q \in \{1, \ldots, n\}\}$ we must be able to find fuzzy sets P and P so that the evaluation of the sentence is not in $\{0, 1\}$.

In [37] it was shown, that method ZS fulfills all the above mentioned properties.

III. Paradox - examples

Consider a set of objects $\{o_1, o_2, \ldots, o_n\}$ that have properties, $f_1, f_2, \ldots, f_m$. Those properties can be described with linguistic values, e.g., for feature f_1, we have $lv_{1,1}, lv_{1,2}, \ldots$. Hence, for each object we can calculate the membership degree to which object o_j is $lv_{k,l}$ for feature f_k.

Without loss of generality summarizer P is $lv_{1,1}$ and qualifier R is $lv_{2,1}$. We consider the form of linguistic summary, "*Most* R objects are P."

We also look at the dual linguistic summary [38], i.e., "$\widehat{most}$ R objects are $\bar{P}$," where $\widehat{most}$ denotes antonym of *most* and $\bar{P}$ denotes complement. The summary and the dual summary should have equal truth value. Without this symmetry we may arrive at a set of inconsistent summaries.

Consider a family of sets with the following characteristics: $O = \{o_i :, i = 1, \ldots, n, : \mu_P(o_i) \geq \mu_R(o_i)\}$. We observe a strange behavior, that is, a paradox when $\max \mu_R(o_i)$ is getting smaller and tends to a value of 0.

As an illustrative example, consider a set of puppies $\{p_1, p_2, ..., p_5\}$. These puppies can be characterized by the properties such as size, fur quality, and appearance. In this example size can be described as *small*, *medium*, *big*. Quality of the fur will be described with linguistic values *soft*, *coarse*, or *rough*. Appearance will be described as either *cute* or *ugly* (not cute). We consider four sets of five puppies, for which we have given their degree of softness and cuteness. Those sets are shown in Tables I-V. ε from set 4, shown in Table IV, is an extremely small value, yet greater than 0.

TABLE I
SET 1

puppy #	softness	cuteness
1	1	1
2	0.9	0.9
3	0	1
4	0	0.1
5	0	0

TABLE II
SET 2

puppy #	softness	cuteness
1	0.5	1
2	0.5	0.9
3	0	1
4	0	0.1
5	0	0

TABLE III
SET 3

puppy #	softness	cuteness
1	0.1	1
2	0.1	0.9
3	0	1
4	0	0.1
5	0	0

TABLE IV
SET 4

puppy #	softness	cuteness
1	ε	1
2	ε	0.9
3	0	1
4	0	0.1
5	0	0

TABLE V
SET 5

puppy #	softness	cuteness
1	0	1
2	0	0.9
3	0	1
4	0	0.1
5	0	0

The reader may notice that the degree of cuteness is the same for the five sets and only the degree of softness varies. Three of the puppies have zero membership for *soft* fur, while two puppies have a degree of membership that gets smaller for each subsequent set. We are interested in the truth value of the linguistic summary "*most soft* puppies are *cute*" and its dual summary "$\widehat{most}$ *soft* puppies are *not cute*.", as this degree of membership falls toward 0.

We used five methods for calculating the *truth value*: two methods based on Zadeh's calculus of quantified propositions (with min and Łukasiewicz as t-norms), the Sugeno integral aggregation [20], and the ZS and GD methods proposed by Delgado, Sanchez, and Vila in [37]. Note that in early work [14], [16] calculations of truth value using Zadeh's calculus assumed only monotonic non-decreasing quantifiers. Problems with this assumption were highlighted in [24], [38], and following these authors we ignore this assumption.

Quantifier *most* is defined as a trapezoidal membership function Trap[0.6, 0.8, 1, 1] and can be calculated as $\max\{\min\{5x - 3, 1\}, 0\}$. Quantifier $\widehat{most}$ is antonym of *most*, i.e., Trap[0,0,0.2,0,4] and can be calculated as $\max\{\min\{-5x + 2, 1\}, 0\}$.

Truth value results are shown in Table VI. We now use these results to illustrate a few issues with the quality measures. First, only one method (the Sugeno integral) managed to obtain equal truth values for the summary and its dual summary across all sets of puppies. All other methods failed to obtain equal truth values for the summary and its dual for both Sets 3 and 4. Those sets contain some puppies that have very small non-zero membership in softness. The GD method also failed to obtain equal truth values for the summary and its dual on

TABLE VI
TRUTH VALUES CALCULATED WITH DIFFERENT EVALUATION METHODS FOR THE FIVE SETS OF DIFFERENT PUPPIES

Set #		Zadeh's calculus (min)	Zadeh's calculus (Lukasiewicz)	Sugeno-based [20]	ZS [37]	GD [37]
1	summary	1	1	1	1	1
1	dual summary	1	1	1	1	1
2	summary	1	1	0.5	1	1
2	dual summary	1	1	0.5	1	0.2
3	summary	1	0	0.1	1	1
3	dual summary	0	1	0.1	0	0
4	summary	1	0	ε	1	1
4	dual summary	0	1	ε	0	0
5	summary	NaN	NaN	0	0	0
5	dual summary	NaN	NaN	0	0	0

Set 2.

Moreover, compare the truth values obtained for the summaries calculated for Sets 4 and 5. Firstly, Zadeh's calculus fails for Set 5, and the choice of t-norm determines the result on Set 4. For the ZS and GD methods the results are more concerning. For small, arbitrary $\epsilon > 0$ the truth values of "Most soft puppies are cute" is 1, but changes to 0 for $\epsilon = 0$. This is a significant change in truth value for a very small change in the data. This change is also contradictory with human intuition, which would suggest that the truth value should approach 0 smoothly as $\epsilon 0$. Furthermore, this step change in truth value at this limit violates one of the required properties for an evaluation method for linguistic summaries, namely that the evaluation must not be too *strict*.

Set 4 seems to support two possible interpretations. Either "*most soft* puppies are *cute*":

1) is true to a very small degree because there is a small degree of support for this in the data (which follows the way of reasoning applied in a rule bases); or,
2) is entirely true as those puppies with a small degree of softness are to a high degree cute.

The ZS and GD methods give a result corresponding to (2), however this seems to go against intuition regarding this data, which says that we don't know much about the cuteness of puppies that are to a high degree also soft, and hence cannot be confident in the statement "*most soft* puppies are *cute*". In other words, the data suggests that "*most* (all) *cute* puppies are *not soft*", which seems at odds with "Most soft puppies are cute".

Unfortunately the Sugeno integral is not the complete solution to avoiding paradoxical truth values for linguistic summaries. Consider the set of puppies shown in Table VII. For this set the summary "*most soft* puppies are *cute*" has a truth of 1, which is also contradictory to human intuition. Clearly, four of the soft puppies – those with 0.99 membership in being soft – are not cute at all, having a membership of merely 0.01 in *cute*.

IV. CONCLUDING REMARKS

We have shown herein a paradox within extended protoform linguistic summaries in the case of objects with low qualifier membership degree for many of the evaluation methods

TABLE VII
SET 6

puppy #	softness	cuteness
1	1	1
2	0.99	0.01
3	0.99	0.01
4	0.99	0.01
5	0.99	0.01

presented. The paradox is clearly visible when we compare the truth values of a summary and its antonym. Only the method based on the Sugeno integral [20] is able to fulfill the requirement of equality of these truth values.

One might argue that in real applications such summaries are never interesting to the user because of low membership values of the qualifier, hence low values of degree of focus [39]. However, given the desirability of producing linguistic summaries from real data sets, it is important to understand how the selected methods will handle these corner cases, and avoid results that violate human intuition (without good cause).

In our future work we will work towards improving the evaluation method based on Sugeno integral, so that inconsistent results will not occur.

ACKNOWLEDGMENT

No puppies were harmed in this research, although some humans were unhappy with puppies being described as ugly. We would like to emphasize that while not all puppies may be cute, or soft, or even puppies anymore, they all deserve a loving home; so, please think of your local animal shelter or pet rescue the next time you're looking for a new animal companion.

REFERENCES

[1] D. Reinsel, J. Gantz, and J. Rydning, "The digitization of the world: From edge to core," 2018. [Online]. Available: https://www.seagate.com/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf

[2] E. Reiter, "Non-experts struggle with information graphics," https://ehudreiter.com/2017/10/02/non-experts-struggle-graphs, 2017.

[3] D. Dubois and H. Prade, "Gradual rules in approximate reasoning," *Information Sciences*, vol. 61, pp. 103–122, 1992.

[4] D. Dubois, H. Prade, and E. Rannou, "User-driven summarization of data based on gradual rules," in *Proceedings of the Sixth IEEE International Conference on Fuzzy Systems*, vol. 2, 1997, pp. 839–844.

[5] A. Oudni, M. Lesot, and M. Rifqi, "Processing contradiction in gradual itemset extraction," in *FUZZ-IEEE 2013, IEEE International Conference on Fuzzy Systems, Hyderabad, India, 7-10 July, 2013, Proceedings.*, 2013, pp. 1–8.

[6] ——, "Accelerating effect of attribute variations: Accelerated gradual itemsets extraction," in *Information Processing and Management of Uncertainty in Knowledge-Based Systems - 15th International Conference, IPMU 2014, Montpellier, France, July 15-19, 2014, Proceedings, Part II*, 2014, pp. 395–404.

[7] A. Wilbik and U. Kaymak, "Gradual linguistic summaries," in *Proceedings of Information Processing and Management of Uncertainty in Knowledge-Based Systems, IPMU 2014,Part II*, 2014, pp. 405–413.

[8] D. Rasmussen and R. R. Yager, "Finding fuzzy and gradual functional dependencies with SummarySQL," *Fuzzy Sets and Systems*, vol. 106, pp. 131–142, 1999.

[9] P. Bosc, D. Dubois, O. Pivet, H. Prade, and M. D. Calmes, "Fuzzy summarization of data using fuzzy cardinalities," in *Proceedings of IPMU 2002 Conference*, 2002, pp. 1553–1559.

[10] G. Raschia and N. Mouaddib, "SAINTETIQ: a fuzzy set-based approach to database summarization," *Fuzzy Sets and Systems*, vol. 129, pp. 137–162, 2002.

[11] M. Ros, M. Pegalajar, M. Delgado, A. Vila, D. T. Anderson, J. M. Keller, and M. Popescu, "Linguistic summarization of long-term trends for understanding change in human behavior," in *Proceedings of the IEEE International Conference on Fuzzy Systems, FUZZ-IEEE 2011*, 2011, pp. 2080–2087.

[12] R. R. Yager, "A new approach to the summarization of data," *Information Sciences*, vol. 28, pp. 69–86, 1982.

[13] J. Kacprzyk, "Intelligent data analysis via linguistic data summaries: a fuzzy logic approach," in *Classification and Information Processing at the Turn of Millennium*, R. Decker and W. Gaul, Eds. Springer-Verlag, Berlin, Heidelberg, New York, 2000, pp. 153–161.

[14] J. Kacprzyk and R. R. Yager, "Linguistic summaries of data using fuzzy logic," *International Journal of General Systems*, vol. 30, pp. 33–154, 2001.

[15] J. Kacprzyk, R. R. Yager, and S. Zadrożny, "A fuzzy logic based approach to linguistic summaries of databases," *International Journal of Applied Mathematics and Computer Science*, vol. 10, pp. 813–834, 2000.

[16] J. Kacprzyk and S. Zadrożny, "Linguistic database summaries and their protoforms: toward natural language based knowledge discovery tools," *Information Sciences*, vol. 173, pp. 281–304, 2005.

[17] ——, "Fuzzy linguistic data summaries as a human consistent, user adaptable solution to data mining," in *Do Smart Adaptive Systems Exist?*, B. Gabrys, K. Leiviska, and J. Strackeljan, Eds. Berlin, Heidelberg, New York: Springer, 2005, pp. 321–339.

[18] R. Castillo-Ortega, N. Marín, and D. Sánchez, "A fuzzy approach to the linguistic summarization of time series," *Multiple-Valued Logic and Soft Computing*, vol. 17, no. 2-3, pp. 157–182, 2011.

[19] G. Smits, P. Nerzic, O. Pivert, and M. Lesot, "Efficient generation of reliable estimated linguistic summaries," in *2018 IEEE International Conference on Fuzzy Systems, FUZZ-IEEE 2018, Rio de Janeiro, Brazil, July 8-13, 2018*, 2018, pp. 1–8.

[20] A. Jain and J. M. Keller, "On the computation of semantically ordered truth values of linguistic protoform summaries," in *2015 IEEE International Conference on Fuzzy Systems, FUZZ-IEEE 2015, Istanbul, Turkey, August 2-5, 2015*, 2015, pp. 1–8.

[21] G. Moyse, M.-J. Lesot, and B. Bouchon-Meunier, "Linguistic summaries for periodicity detection based on mathematical morphology," in *2013 IEEE Symposium on Foundations of Computational Intelligence (FOCI)*, 2013, pp. 106–113.

[25] P. S. Szczepaniak and J. Ochelska, "Linguistic summaries of standardized documents," in *Advances in Web Intelligence and Data Mining*,

[22] J. Kacprzyk, A. Wilbik, and S. Zadrożny, "Linguistic summarization of time series using a fuzzy quantifier driven aggregation," *Fuzzy Sets and Systems*, vol. 159, no. 12, pp. 1485–1499, 2008.

[23] ——, "An approach to the linguistic summarization of time series using a fuzzy quantifier driven aggregation," *International Journal of Intelligent Systems*, vol. 25, no. 5, pp. 411 – 439, 2010.

[24] A. Wilbik, J. M. Keller, and G. L. Alexander, "Linguistic summarization of sensor data for eldercare," in *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics (SMC 2011)*, 2011, pp. 2595–2599.

M. Last, P. S. Szczepaniak, Z. Volkovich, and A. Kandel, Eds. Springer Berlin Heidelberg, 2006, pp. 221–232.

[26] D. Anderson, R. H. Luke, J. M. Keller, M. Skubic, M. Rantz, and M. Aud, "Linguistic summarization of video for fall detection using voxel person and fuzzy logic," *Computer Vision and Image Understanding*, vol. 1, no. 113, pp. 80–89, 2009.

[27] ——, "Modeling human activity from voxel person using fuzzy logic," *IEEE Transactions on Fuzzy Systems*, vol. 1, no. 17, pp. 39–49, 2009.

[28] D. Anderson, R. H. Luke, E. Stone, and J. M. Keller, "Segmentation and linguistic summarization of voxel environments using stereo vision and genetic algorithms," in *Proceedings IEEE International Conference on Fuzzy Systems, World Congress on Computational Intelligence*, 2010, pp. 2756–2763.

[29] A. Wilbik and R. Dijkman, "Linguistic summaries of process data," in *Proceedings of IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2015)*, 2015.

[30] R. M. Dijkman and A. Wilbik, "Linguistic summarization of event logs - A practical approach," *Inf. Syst.*, vol. 67, pp. 114–125, 2017.

[31] M. Delgado, M. D. Ruiz, D. Sanchez, and M. A. Vila, "Fuzzy quantification: a state of the art," *Fuzzy Sets and Systems*, vol. 242, pp. 1 – 30, 2014.

[32] J. Kacprzyk and P. Strykowski, "Linguistic summaries of sales data at a computer retailer: a case study," in *Proceedings of IFSA'99*, vol. 1, 1999, pp. 29–33.

[33] A. Bugarín, N. Marín, D. Sánchez, and G. Trivino, "Aspects of quality evaluation in linguistic descriptions of data," in *2015 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, 2015, pp. 1–8.

[34] N. Marín and D. Sánchez, "On generating linguistic descriptions of time series," *Fuzzy Sets and Systems*, vol. 285, pp. 6 – 30, 2016, special Issue on Linguistic Description of Time Series.

[35] L. A. Zadeh, "Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic," *Fuzzy Sets and Systems*, vol. 9, no. 2, pp. 111–127, 1983.

[36] J. Kacprzyk, A. Wilbik, and S. Zadrożny, "Linguistic summarization of time series under different granulation of describing features," in *Rough Sets and Intelligent Systems Paradigms - RSEISP 2007*, M. Kryszkiewicz, J. F. Peters, H. Rybinski, and A. Skowron, Eds. Springer-Verlag, Berlin and Heidelberg, 2007, pp. 230–240.

[37] M. Delgado, M. J. Martín-Bautista, D. Sánchez, and M. A. Vila, "Fuzzy cardinality based evaluation of quantified sentences," *International Journal of Approximate Reasoning*, vol. 23, pp. 23–66, 2000.

[38] G. Moyse, M. Lesot, and B. Bouchon-Meunier, "Oppositions in fuzzy linguistic summaries," in *2015 IEEE International Conference on Fuzzy Systems, FUZZ-IEEE 2015, Istanbul, Turkey, August 2-5, 2015*, 2015, pp. 1–8.

[39] J. Kacprzyk and A. Wilbik, "Towards an efficient generation of linguistic summaries of time series using a degree of focus," in *Proceedings of the 28th North American Fuzzy Information Processing Society Annual Conference – NAFIPS 2009*, 2009.