Linguistic Interpretation of Natural Data using New Forms of Intermediate Quantifiers

MURINOVÁ Petra[1] and FIALA Karel[1]

[1] *University of Ostrava, Institute for Research and Applications of Fuzzy Modeling, 30. dubna 22, 701 03 Ostrava 1, Czech Republic*
E-mail: {petra.murinova,karel.fiala}@osu.cz

Abstract

This paper examines the application of fuzzy natural logic in the analysis of scientific data and their representation through special linguistic expressions. We use the theory of evaluative linguistic expressions, which make it possible to describe quantitative data using imprecise espressions such as “very small”, “medium”, “large”, and similar. They occur in the definition of the, so called, intermediate quantifiers using which we characterize given data.

Keywords: Fuzzy natural logic, Intermediate quantifiers, Linguistic association, Graded cube of opposition

1 Introduction

Generalized quantifiers have proven effective in providing linguistic summaries of numerical data.The latter is a part of the data-to-text data-mining paradigm, which facilitates the extraction of knowledge from large numerical datasets in the form of fuzzy association rules or special summarising expressions ([1, 2]). The main aim of this paper is to build on theoretical insights into intermediate quantifiers presented in the earlier works (see [3, 4, 5]) and to propose a linguistic description of the real-world data.

Below are examples of natural language expressions with quantifiers related to graded Peterson’s cube of opposition (see [6]).

- Most people who do not eat enough vegetables are more likely to suffer from digestive problems.
- Most people who are not physically active tend to gain weight.
- Almost all people who do not get enough sleep have trouble concentrating the next day.

2 Preliminaries

2.1 Mathematical bacground

Intermediate quantifiers are formally defined within a formal theory T^{IQ} of Łukasiewicz Fuzzy Type Theory (Ł-FTT) determined by 15 special axioms. Its underlying algebra of truth values is an MV$_\Delta$-algebra $\mathscr{E} = \langle E, \vee, \wedge, \otimes, \rightarrow, \mathbf{0}, \mathbf{1}, \Delta \rangle$ where $\mathbf{0}$ and $\mathbf{1}$ are minimal and maximal elements, respectively. We usually assume the standard Łukasiewicz MV$_\Delta$-algebra, where the domain of truth degrees is given by the interval $E = [0, 1]$.

Acknowledgement The paper is supported by the project “Research of Excellence on Digital Technologies and Wellbeing CZ.02.01.01/00/22_008/0004583” which is co-financed by the European Union.

https://www.doi.org/10.15452/978-80-7599-515-5.2026.15

2.2 Definition of intermediate quantifiers

Below we present mathematical definitions of fuzzy intermediate quantifiers that define the back face graded Peterson's cube of opposition. For the detail see [6].

Definition 1 (Intermediate quantifiers related to cube of opposition) *Let $Ev \in Form_{oo}$ be a formula of type oo representing an evaluative expression, $x \in Form_{\alpha}$ be a variable and $A, B, z \in Form_{\alpha}$ be formulas where α is a selected type. Then either of the formulas*

$$(Q^{\forall}_{Ev}\, x)(\neg B, \neg A) \equiv (\exists z)[(\forall x)((\neg B|z)\, x \Rightarrow \neg Ax) \wedge Ev((\mu(\neg B))(\neg B|z))], \quad (1)$$

$$(Q^{\exists}_{Ev}\, x)(\neg B, \neg A) \equiv (\exists z)[(\exists x)((\neg B|z)x \wedge \neg Ax) \wedge Ev((\mu(\neg B))(\neg B|z))]. \quad (2)$$

construes the sentence "⟨Quantifier⟩ not B's are not A".

Specific quantifier is determined by a concrete evaluative expression. For example, "Almost all" which is denoted by **P** is determined by Ev ="extremely big". For further quantifiers, see, e.g. [5].

3 Application

In this section, we present the application of the theory of intermediate quantifiers to the penguins dataset containing measurements of adult foraging penguins near Palmer Station [7], available in the R package "palmerpenguins" [8]. It contains 344 observations and 8 variables. We removed 11 observations containing missing (NA) values. We chose the following real valued variables:

- bill_length (millimeters),
- bill_depth (millimeters),
- flipper_length (millimeters),
- body_mass (grams).

We used R programming language [9] along with the "lfl" package [10]. First, we transformed values of the variables into membership degrees of fuzzy sets. For this, we applied the theory of evaluative linguistic expressions [11]. For simplicity, we considered the basic evaluative trichotomy "small", "medium", and "big" only. Then we expressed these fuzzy sets using intermediate quantifiers. Examples of the results are presented in Table 1.

Table 1: Examples of degrees of truth of various quantifiers

	Quantifier					
Implication	**all**	**almost all**	**most**	**many**	**several**	**a few**
Small bill depth ⇒ medium body mass	0.24	0.48	0.6	1	1	1
Medium bill depth ⇒ Non big flipper length	0	0.32	0.6	1	1	1
Non small flipper length⇒ big bill length	0	0	0	0.06	0.13	0.2
Non big body mass ⇒ non big bill depth	0	0.58	0.77	1	1	1

Table 1 can be used for derivation of linguistic expressions characterizing the given data. For example:

- "**Most (T)** penguins who do not have a big body mass also do not have a big bill depth." (truth value 0.77). When considering the highest truth value, we can replace the quantifier **Most (T)** by **Many (K), Several (S), A few (F)** (truth value 1).
- "**Almost all (P)** penguins who have a medium bill depth do not have a big flipper length" (truth value 0.32). Similarly, the highest truth value 1 has the quantifier **Many (K), Several (S), A few (F)**.
- "**Several (S)** penguins who do not have a small flipper length have a big bill length" (truth value 0.13).

Using the theory of logical inference, we can derive additional insight by applying selected forms of valid logical syllogisms. The following is an example of a syllogism related to the graded Peterson's cube of opposition:

$$\mathbf{dAO-IV}: \begin{array}{l} \mathscr{P}_1 : \textit{Many penguins which have not small flipper length have big bill depth.} \\ \mathscr{P}_2 : \textit{All small bill depth have medium body mass.} \\ \hline \mathscr{C} : \textit{Some penguins with small body mass have not small flipper length.} \end{array}$$

Syntactic proofs of the validity of similar syllogisms for specific cases can be found in [12]. Much richer information can be obtained when using linguistic hedges, e.g. "very, extremely, more or less, roughly", etc.

Since we are limited by the number of pages, we refer readers to the cited publications for more detailed results.

4 Conclusion

In this paper, we focused on interpreting data using generalized intermediate quantifiers related to the graded Peterson's cube of opposition. We gave examples of linguistic expressions interpreting the data. We also outlined that another related area is the theory of logical inference, which allows us to derive new conclusions.

References

[1] R. R. Yager, A new approach to the summarization of data, Information Sciences 28 (1) (1982) 69–86.

[2] J. Kacprzyk, R. R. Yager, S. Zadrożny, A fuzzy logic based approach to linguistic summaries of databases, International Journal of Applied Mathematics and Computer Science 10 (4) (2000) 813–834.

[3] P. Murinová, V. Novák, A formal theory of generalized intermediate syllogisms, Fuzzy Sets and Systems 186 (2013) 47–80.

[4] S. Boffa, P. Murinová, V. Novák, P. Ferbas, Graded cubes of opposition in fuzzy formal concept analysis, International Journal and Aproximate Reasoning 145 (2022) 187–209.

[5] P. Murinová, V. Novák, The theory of intermediate quantifiers in fuzzy natural logic revisited and the model of "many", Fuzzy sets and systems 388 (2020) 56–89.

[6] P. Murinová, Graded structures of opposition in fuzzy natural logic, Logica Universalis 265 (2020) 495–522.

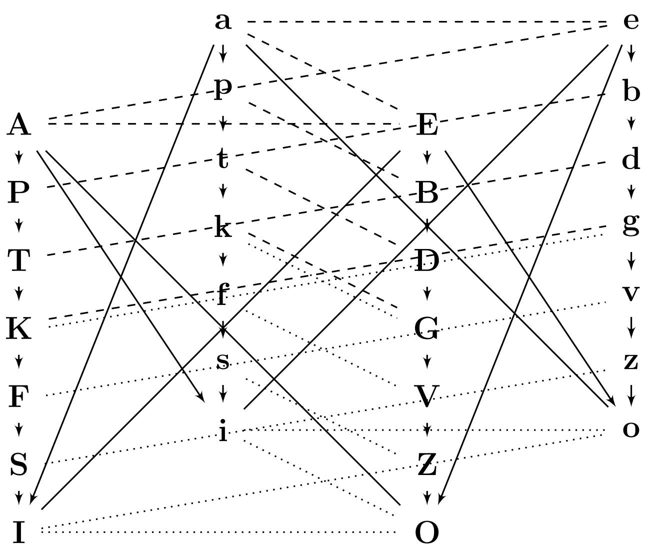

Figure 1: Graded Peterson's cube of opposition. The letters refer to specific intermediate quantifiers (see [12]).

[7] K. B. Gorman, T. D. Williams, W. R. Fraser, Ecological sexual dimorphism and environmental variability within a community of antarctic penguins (genus pygoscelis), PloS one 9 (3) (2014) e90081.

[8] A. M. Horst, A. P. Hill, K. B. Gorman, palmerpenguins: Palmer Archipelago (Antarctica) penguin data, r package version 0.1.0 (2020). doi:10.5281/zenodo.3960218.
URL https://allisonhorst.github.io/palmerpenguins/

[9] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria (2025).
URL https://www.R-project.org/

[10] M. Burda, Linguistic fuzzy logic in R, in: Proc. IEEE International Conference On Fuzzy Systems (FUZZ-IEEE), IEEE, Istanbul, 2015, pp. 1–7.

[11] V. Novák, A comprehensive theory of trichotomous evaluative linguistic expressions, Fuzzy Sets and Systems 159 (22) (2008) 2939–2969.

[12] K. Fiala, P. Murinová, A formal analysis of generalized Peterson's syllogisms related to graded Peterson's cube, Mathematics 10(6) (2022).