

From Zadeh's Computing with Words towards eXplainable Artificial Intelligence

Jose M. Alonso¹

Key words: Fuzzy Logic, Explainable AI, Computing with Words, Computing with Perceptions, Cointension, Interpretable Fuzzy Systems

Extended Abstract

The European Commission has identified Artificial Intelligence (AI) as the “most strategic technology of the 21st century” [7]. AI is already part of our everyday life through many successful applications into real-world usage and according to Accenture [16] the economic impact of the automation of knowledge work, robots and self-driving vehicles could reach between 6.5 and 12 €trillion annually by 2025. People are used to buzzwords like smart watch, smart phone, smart home, smart car, smart city, etc. In practice, we are surrounded by smart gadgets, i.e., devices connected to Internet and endowed with some level of autonomy and intelligence thanks to AI systems. The cohabitation of humans and smart gadgets makes society demand the development of a new generation of explainable AI systems, i.e., AI systems ready to explain naturally (as humans do) their automatic decisions.

Thus, the research field on explainable AI is flourishing and attracting more and more attention not only regarding technical but also ethical and legal issues [8]. The ACM Code of Ethics [1] highlighted explanation as a basic principle in the search for “Algorithmic Transparency and Accountability”. In addition, Floridi et al. defined the concept of “explicability” in reference to both “intelligibility” and “explainability” and hence captured the need for transparency and for accountability in an ethical framework for AI [10]. Moreover, the new European General Data Protection Regulation (GDPR) [14] refers to the “right to explanation”, i.e., GDPR states that European citizens have the right to ask for explanations of decisions affecting them, no matter who (or what AI system) makes such decision.

The term eXplainable Artificial Intelligence (XAI) was coined by the USA Defense Advanced Research Projects Agency (DARPA) [11]. Assuming that “*even though current AI systems offer many benefits in many applications, their effectiveness is limited by a lack of explanation ability when interacting with humans*” DARPA launched to the research community (including both academy and industry) the challenge of designing new self-explanatory AI systems from 2017 to 2021.

In Europe, there is not any initiative similar to the DARPA challenge on XAI yet. However, the European Commission has already pointed out the convenience of launching a pilot in XAI [7]. In June 2018, the Confederation of Laboratories for Artificial Intelligence Research in Europe (CLAIRE^a), a novel initiative to create a network of excellence in AI with the most well-recognized universities and R+D centres, emphasized in its European vision for AI the need to search for transparent, explainable, fair and socially compatible intelligent systems. Moreover, The AI4EU^b

Email addresses: josemaria.alonso.moral@usc.es (Jose M. Alonso)

¹Centro Singular de Investigación en Tecnoloxías da Información (CiTIUS), Universidade de Santiago de Compostela, Campus Vida, E-15782, Santiago de Compostela, Galicia, Spain

Research Centre in Information Technologies (CiTIUS), University of Santiago de Compostela (USC) (May 2018).

Final version published in Proceedings of the 12th International Workshop on Fuzzy Logic and Applications (WILF2018).

Please cite this article in press as:

J.M. Alonso, “From Zadeh's computing with words towards explainable Artificial Intelligence”, *WILF2018 - 12th International Workshop on Fuzzy Logic and Applications*, Springer, pp. 244-248, 2019, DOI: 10.1007/978-3-030-12544-8_21

Available online at http://dx.doi.org/10.1007/978-3-030-12544-8_21

^a <https://claire-ai.org/>

^b <http://ai4eu.org/>

H2020 Project is funded by call ICT-26 2018 (grant 825619) with the aim of: (1) to mobilize the entire European AI community to make AI promises real for the European Society and Economy; and (2) to create a leading collaborative AI European platform to nurture economic growth. Explainable Human-centered AI is highlighted as one of the five key research areas to consider and it is present in 5 out of the 8 experimental pilots to be developed.

In the rest of this manuscript we briefly review a selection of outstanding Zadeh's contributions which are likely to have direct impact in the research field of XAI. The paradigm of Computing with Words (CWW) is especially relevant because humans are used to explanations in natural language (NL).

From Prof. Zadeh's seminal ideas on fuzzy sets and systems [21], many key concepts such as linguistic variables and linguistic rules have turned up in the field of Fuzzy Logic (FL). Accordingly, FL has many successful applications [19]. In addition, as it is described in [4], about 30% of publications in XAI come from authors well recognized in the field of FL. This is mainly due to the commitment of the fuzzy community to produce interpretable fuzzy systems [3]. Actually, interpretability is deeply rooted in the fundamentals of FL. However, it is worthy to note that interpretability is not guaranteed only because of applying FL. In practice, producing interpretable fuzzy systems is a matter of careful design [17].

In XAI, interpretability is a key issue but understandability and comprehensibility which are not so deeply considered by the FL community also play a prominent role. Nowadays, a new generation of intelligent systems is expected to provide users with natural explanations. Those explanations should be easy to understand no matter the user background. Since, humans think and compute naturally with words, explanations in NL are likely to be considered as natural explanations. Prof. Zadeh was the first to talk about CWW [22] as an extension of fuzzy sets and systems. Later, Prof. Kacprzyk gave some hints about how to implement CWW [12]. Moreover, he highlighted the need to connect CWW with the paradigm of NL Generation (NLG) [9]. It is worth noting that NLG is a well-known area within the Computational Linguistics and AI research fields. The connection between FL and NLG has been further researched by other authors [2, 15].

In addition, Prof. Zadeh was pioneer to introduce a new generation of more natural intelligent systems, ready to compute with perceptions and make approximate reasoning as humans naturally do. Thus, the Computational Theory of Perceptions (CTP) [20, 23] was first introduced by Zadeh and later applied by Trivino and Sugeno to automatically generate linguistic descriptions of complex phenomena [18]. The CTP has been successfully applied for example to explain the energy consumption at home [5] or to automatically generate linguistic descriptions associated to the USA census data [6].

In addition, Prof. Zadeh also coined the concept of cointension [24]. The semantic-cointension approach [13] is already applied to assess interpretability of fuzzy systems. Likewise, it can be considered when evaluating the understandability of explanations in XAI. In short, two different concepts referring almost to the same entities are taken as cointensive. Accordingly, an explanation in NL is deemed as comprehensible only when the explicit semantics embedded in it is cointensive with the implicit semantics inferred by the user when reading and processing the given explanation.

To sum up, Prof. Zadeh made many highly valuable contributions to the FL field and beyond. Many of these contributions were pioneer ideas and/or challenging proposals with a lot of potential to be fully developed later by other researchers. Nowadays, XAI is a prominent and fruitful research field where many of Zadeh's contributions can become crucial if they are carefully considered and thoroughly developed. For example, two major open challenges for XAI are: (1) how to build conversational agents able to provide humans with semantic grounding, persuasive and trustworthy interactive explanations; and (2) how to measure the effectiveness and naturalness of automatically generated explanations. CWW as well as fuzzy measures and Z-numbers [25] introduced by Zadeh are likely to contribute to successfully address both challenges and achieve valuable results.

Acknowledgments

Jose M. Alonso is *Ramón y Cajal* Researcher (RYC-2016-19802). In addition, this research was funded by the Spanish Ministry of Economy and Competitiveness (grants TIN2014-56633-C3-1-R, TIN2017-90773-REDT and TIN2017-84796-C2-1-R) and the Galician Ministry of Education (grants ED431F 2018/02, GRC2014/030 and "accreditation 2016-2019, ED431G/08"). All grants were co-funded by the European Regional Development Fund (ERDF/FEDER program).

References

- [1] ACM US Public Policy Council: Statement on Algorithmic Transparency and Accountability (2017), http://www.acm.org/binaries/content/assets/public-policy/2017{_}usacm{_}statement{_}algorithms.pdf
- [2] Alonso, J.M., Bugarin, A., Reiter, E.: Special issue on natural language generation with computational intelligence. *IEEE Computational Intelligence Magazine* **12**(3), 8–9 (2017). <https://doi.org/10.1109/MCI.2017.2708919>
- [3] Alonso, J.M., Castiello, C., Mencar, C.: Interpretability of fuzzy systems: Current research trends and prospects. In: *Handbook of Computational Intelligence*, pp. 219–237. Springer (2015)
- [4] Alonso, J.M., Castiello, C., Mencar, C.: A bibliometric analysis of the explainable artificial intelligence research field. In: *17th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU)*. vol. CCIS853, pp. 2–15 (2018)
- [5] Conde-Clemente, P., Alonso, J.M., Trivino, G.: Towards automatic generation of linguistic advice for saving energy at home. *Soft Computing* **22**(2), 345–359 (2018)
- [6] Conde-Clemente, P., Trivino, G., Alonso, J.M.: Generating Automatic Linguistic Descriptions with Big Data. *Inf. Sciences* **380**(2), 12–30 (2017)
- [7] European Commission: Artificial Intelligence for Europe. Tech. Rep. (2018), <https://ec.europa.eu/digital-single-market/en/news/communication-artificial-intelligence-europe>
- [8] EU AI HLEG: AI Ethics Guidelines. Tech. Rep. (2019), <https://ec.europa.eu/digital-single-market/en/news/draft-ethics-guidelines-trustworthy-ai>
- [9] Gatt, A., Krahmer, E.: Survey of the state of the art in natural language generation: Core tasks, applications and evaluation. *Journal of Artificial Intelligence Research* **61**, 65–170 (2018)
- [10] Floridi, L. et al.: AI4People - An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. *Minds and Machines* **28**(4), 689–707 (2018)
- [11] Gunning, D.: Explainable Artificial Intelligence (XAI). Tech. rep., Defense Advanced Research Projects Agency (DARPA), Arlington, USA (2016), DARPA-BAA-16-53, <https://www.darpa.mil/program/explainable-artificial-intelligence>
- [12] Kacprzyk, J.: Computing with words is an implementable paradigm: Fuzzy queries, linguistic data summaries, and natural-language generation. *IEEE Transactions on Fuzzy Systems* **8**, 451–472 (2010)
- [13] Mencar, C., Castiello, C., Cannone, R., Fanelli, A.M.: Interpretability assessment of fuzzy knowledge bases: A cointension based approach. *International Journal of Approximate Reasoning* **52**(4), 501–518 (2011)
- [14] Parliament and Council of the European Union: General data protection regulation (GDPR) (2016), <http://data.europa.eu/eli/reg/2016/679/oj>
- [15] Ramos-Soto, A., Bugarin, A., Barro, S.: On the role of linguistic descriptions of data in the building of natural language generation systems. *Fuzzy Sets and Systems* **285**, 31–51 (2016)
- [16] Schoeman, W.: Why AI is the future of growth. Accenture Tech. Rep. (2016), <https://www.accenture.com/za-en/company-news-release-why-artificial-intelligence-future-growth>
- [17] Trillas, E., Eciolaza, L.: *Fuzzy Logic: An Introductory Course for Engineering Students*. Springer (2015)
- [18] Trivino, G., Sugeno, M.: Towards linguistic descriptions of phenomena. *International Journal of Approximate Reasoning* **54**, 22–34 (2013)
- [19] Yager, R.R., Zadeh, L.A.: *An Introduction to Fuzzy Logic Applications in Intelligent Systems*. The Springer International Series in Engineering and Computer Science, Springer US (2012)
- [20] Zadeh, L.A.: Toward a perception-based theory of probabilistic reasoning with imprecise probabilities. *Statistical Planning and Inference* **105**, 233–264 (2002)
- [21] Zadeh, L.A.: Fuzzy sets. *Information and Control* **8**(3), 338–353 (1965)
- [22] Zadeh, L.A.: From computing with numbers to computing with words - From manipulation of measurements to manipulation of perceptions **46**(1), 105–119 (1999)
- [23] Zadeh, L.A.: A new direction in AI: toward a computational theory of perceptions. *Artificial Intelligence Magazine* **22**(1), 73–84 (2001)
- [24] Zadeh, L.A.: Is there a need for fuzzy logic? *Inf. Sciences* **178**, 2751–2779 (2008)
- [25] Zadeh, L.A.: A note on z-numbers *Inf. Sciences* **181**, 2923–2932 (2011)