All We (and LLMs) Need Is Fuzzy: An Argument

Olga Kosheleva[1] and Vladik Kreinovich[2]
Departments of [1]Teacher Education and [2]Computer Science
University of Texas at El Paso
500 W. UNiversity
El Paso, TX 79968, USA
olgak@utep.edu. vladik@utep.edu

Abstract

Large Language Models (LLMs) like ChatGPT have spectacular successes – but they also have surprising failures that an average person with common sense could easily avoid. It is therefore desirable to incorporate the imprecise ("fuzzy") common sense into LLMs. A natural question is: to what extent will this help? This way, we may avoid a few simple mistakes, but will it significantly improve the LLMs' performance? What portion of the gap between current LLMs and ideal perfect AI-based agents can be, in principle, covered by using fuzzy techniques? Judging by the fact that few researchers working on LLMs (and on deep learning in general) try fuzzy methods shows that most these researchers do not believe that the use of fuzzy techniques will significantly improve LLMs' performance. Contrary to this pessimistic viewpoint, our analysis shows that potentially, fuzzy techniques can cover all the above gap – or at least a significant portion of this gap. In this sense, indeed, all LLMs need to become perfect is fuzzy techniques.

1 Formulation of the problem

LLMs are great but not perfect. Modern AI techniques, in particular, Large Language Models (LLMs), have achieved so many spectacular successes that many of us have become in awe of them – and our only problem seems to be that they are so smart that they can take over us. In doing this, we forget that while, in general, ChatGPT and LLMs produce impressive results, once in a while they produce results that we humans can easily see as wrong. This is not just complex LLMs:

- several crashes of AI-controlled self-driving cars occurred in traffic situations in which even a not-very-experienced human driver would know how to avoid;

- one of the most spectacular successes of deep-learning-based AI – winning over a human world champion in Go – was recently kind of overturned by a not-very-highly-ranked Go player who beat AI by using rather simple moves, moves that most human players would know how to react to.

So what is missing? In both examples, what is missing is not the ability to deal with complex situations, what is missing is simple common sense. So, to improve the situation, it makes sense to take common sense into account.

A similar challenge happened 60 years ago. How can we take common sense into account? This question was first asked, in the 1960s, by Lotfi Zadeh, one of the leading control experts of that time, who encountered another challenge: that optimized automatic controllers often performed worse than human controllers. The answer to this challenge seemed to be straightforward: incorporate the knowledge of expert controllers into the automatic control systems. However, it was not clear how to follow this recommendation. Many expert controllers were willing to share their strategies, but the problem was that they did not describe these strategies in computer-understandable precise form, they could inly describe their strategies by using imprecise ("fuzzy") words from natural language like "small".

To overcome this challenge, Zadeh came up with a technique – that he called *fuzzy* – for transforming this natural-language description into precise computer-understandable control strategies. This technique indeed led to many successes (see, e.g., [1, 2, 3, 5, 6, 10]) – although, of course, it is not a panacea.

So maybe fuzzy technique can help here as well? So a natural idea is to try to use fuzzy techniques to help LLMs common sense – o, to be more cautious, to acquire more of common sense.

But will this be enough? Probably fuzzy techniques will lead to some successes, but is using these techniques the right research direction? Very few people in the current AI community follow this path. THis means that the vast majority of them do not believe that using fuzzy techniques will drastically improve the situation. And their reasoning seems to make sense: after all, fuzzy successes are mainly in the past, these successes often pale in comparison with successes of modern deep learning techniques.

We arrive at the following research question. In view of the widely spread pessimism about fuzzy, to convince researchers to try fuzzy techniques, it is desirable to estimate how much fuzzy can help.

What we do in this paper. In this paper, we use common sense (pun intended) to provide such an estimate – and our estimate shows that fuzzy techniques have a potential to (almost) close the gap between current LLMs and ideal future common-sense-using AI-based agents.

2 How we solve this problem

What we plan to do in this paper. Our main idea is that LLMs use only crisp – precise – part of the information: namely, they use the facts. The LLMs do not use fuzzy (imprecise) expert knowledge. In order to show that fuzzy techniques have a potential to close the current gap, we need:

- first, to gauge the size of this gap – i.e., to analyze what portion of information is missing, and
- second, to gauge what portion of information is fuzzy.

To perform the second task, we need to recall the main ideas behind fuzzy techniques. Once both tasks are performed, we will be able to compare the portions and thus, to estimate to what extend fuzzy techniques can help.

How far are LLMs from common sense? In order to show that fuzzy techniques have a potential to close this gap, we need to gauge the size of this gap. We want a general estimate, applicable for all kinds of LLMs and AIs, not just one specific model. Because of this desire, we selected the paper [9] that analyzed several different LLMs – on the example of predicting prices of gold and other precious metals.

According to this paper, the correlation between these predictions and real data is about 20% for all the LLMs. One may hope that if we combine different LLMs, the gap will decrease, so that some of the LLMs will pick up where others fail. This would have been the case if the results of these LLMs were independent – then by combining them, we would indeed get more accurate results. Unfortunately, these hopes are in vain: LLMs' results are highly correlated: the correlation between every two of them is about 70-80%.

So, the $100 - 20 = 80\%$ is not just a gap of each LLM, it is a joint gap of all LLMs.

Let us recall the main ideas behind fuzzy techniques. To analyze what part of information is fuzzy, let us briefly recall how fuzzy techniques work. In these techniques, for each imprecise property like "x is small", with each possible value of the quantity x, we associate a degree $\mu(x)$ from the interval $[0,1]$ to which this value x satisfies the given property (e.g., to which x is small). Here:

- the value 1 means that we are absolutely sure that x *has* the given property,
- the value 0 means that we are absolutely sure that x *does not have* the given property, and
- values between 0 and 1 correspond to intermediate degrees of confidence.

The resulting function $\mu(x)$ is known as a *membership function*, or, alternatively, as a *fuzzy set*.

It is well known that to process fuzzy data, it is convenient to use an alternative representation of fuzzy sets – via so-called α-*cuts*, i.e., sets

$$\mathbf{x}(\alpha) \stackrel{\text{def}}{=} \{x : \mu(x) \geq \alpha\}$$

for $\alpha > 0$ and $\mathbf{x}(0) \stackrel{\text{def}}{=} \overline{\{x : \mu(x) > 0\}}$ for $\alpha = 0$.

It is known that once we have all the α-cuts, we can uniquely reconstruct the original membership function.

The meaning of α-cuts is as follows: For each α and for each $x \notin \mathbf{x}(\alpha)$, our degree of confidence that this x is possible is smaller than α. Thus, our degree of confidence that this x is *not* possible is larger than $1 - \alpha$. Thus, with degree of confidence $1 - \alpha$, we are sure that all possible values x are located in the corresponding α-cut $\mathbf{x}(\alpha)$.

Let us estimate which portion of information is stored in non-crisp form. From the purely theoretical viewpoint, to reconstruct the membership function, we need to know α-cuts corresponding to *all* infinitely many values α from the interval $[0, 1]$. The reason for this need is that in principle, a degree $\mu(x)$ can be any value from the interval $[0, 1]$.

However, it is not possible for an expert to meaningfully distinguish between, e.g., degree 0.8 and degree 0.81. According to the psychological seven-plus-minus-two law (see, e.g., [4, 7]), a human being can meaningly distinguish only between 7 ± 2 different values – i.e., at best, between $7 + 2 = 9$ values. We want possible values to include 0 (absolutely false) and 1 (absolutely true). This leaves us with 7 intermediate values. For simplicity, it makes sense to assume that these intermediate values are uniformly spread on the interval $[0, 1]$, i.e., that they have the form 0, $1/8$, $2/8$, ..., $7/8$, and 1. We may have other values of $\mu(x)$, but these values are indistinguishable from these nine ones. So, without losing any expert information, we can safely assume that all the values $\mu(x)$ are equal to one of these nine numbers.

For such membership functions, we do not need to know all infinitely many α-cuts: it is sufficient to use α-cuts corresponding to the above nine values α. Using the above general meaning of α-cuts, we can make the following conclusions:

- we are fully confident that the actual value is in $\mathbf{x}(0)$;
- with confidence $7/8$, we are sure that the actual value is in $\mathbf{x}(1/8)$;
- ...
- with confidence $1 - i/8$, we are sure that the actual value is in $\mathbf{x}(i/8)$;
- ...

So, in this sense, we have nine pieces of information:

- one piece with confidence 1 – this is the crisp (non-fuzzy) piece and

- eight fuzzy pieces with confidences, correspondingly, 7/8, 6/8, ..., and 0.

To get the amount of information contained in each piece, it makes sense to multiply the average amount of knowledge in a corresponding statement by the degree of confidence. For example, if our degree of confidence in a statement is 0, this means that we have no information at all.

Thus, the overall amount of information contained in all 9 pieces is proportional to the sum

$$1+\frac{7}{8}+\frac{6}{8}+\ldots+\frac{1}{8}+0=\frac{8+7+6+\ldots+1+0}{8}=\frac{1}{8}\cdot(1+2+\ldots+8)=$$

$$\frac{1}{8}\cdot\frac{8\cdot(8+1)}{2}=\frac{9}{2}=4.5,$$

while the amount of information contained in the crisp part is proportional to 1. Hence, the proportion of information contained in the fuzzy part is equal to

$$\frac{4.5-1}{4.5}\approx 0.78.$$

Conclusion. What is missing is approximately 80% of the information, and what fuzzy can bring is about 78%. Taking into account that these are crude estimates, we can reasonably conclude that fuzzy information can fill the gap between current LLMs and the ideal AI-based agents – or at least fill in the significant portion of the bill.

Remaining open question. Of course, the big question is how to incorporate fuzzy knowledge into LLMs. We still do not know how to do it.

But the fact that we do not know how to do it does not mean that this paper is useless. Our analysis – presented in this paper – shows that fuzzy knowledge has a potential of filling the gap. So hopefully this will inspire more researchers to work in this direction – the direction of analyzing how to incorporate imprecise (fuzzy) expert knowledge into the LLMs.

Acknowledgments

This work was supported in part by the National Science Foundation grants 1623190 (A Model of Change for Preparing a New Generation for Professional Practice in Computer Science), HRD-1834620 and HRD-2034030 (CAHSI Includes), EAR-2225395 (Center for Collective Impact in Earthquake Science C-CIES), and by the AT&T Fellowship in Information Technology.

It was also supported by a grant from the Hungarian National Research, Development and Innovation Office (NRDI), and by the Institute for Risk and Reliability, Leibniz Universitaet Hannover, Germany.

References

[1] R. Belohlavek, J. W. Dauben, and G. J. Klir, *Fuzzy Logic and Mathematics: A Historical Perspective*, Oxford University Press, New York, 2017.

[2] G. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic*, Prentice Hall, Upper Saddle River, New Jersey, 1995.

[3] J. M. Mendel, *Explainable Uncertain Rule-Based Fuzzy Systems*, Springer, Cham, Switzerland, 2024.

[4] G. A. Miller, "The magical number seven plus or minus two: some limits on our capacity for processing information", *Psychological Review*, 1956, Vol. 63, No. 2, pp. 81–97.

[5] H. T. Nguyen, C. L. Walker, and E. A. Walker, *A First Course in Fuzzy Logic*, Chapman and Hall/CRC, Boca Raton, Florida, 2019.

[6] V. Novák, I. Perfilieva, and J. Močkoř, *Mathematical Principles of Fuzzy Logic*, Kluwer, Boston, Dordrecht, 1999.

[7] S. K. Reed, *Cognition: Theories and Application*, SAGE Publications, Thousand Oaks, California, 2022.

[8] R. Waters, "Man beats machine at Go in human victory over AI: Amateur exploited weakness in systems that have otherwise dominated grandmasters", *Financial Times*, February 19, 2023,
https://arstechnica.com/information-technology/2023/02/man-beats-machine-at-go-in-human-victory-over-ai/

[9] Y. Yoon, "Demystifying Stochastic Gradient Descent and its variants: practical insights and Python implementations", In: V. Kreinovich, W. Yamaka, and S. Leurcharusmee (eds.), *Data Science for Econometrics and Related Topics*, Springer, Cham, Switzerland, to appear.

[10] L. A. Zadeh, "Fuzzy sets", *Information and Control*, 1965, Vol. 8, pp. 338–353.