

Towards the Most Robust Way of Assigning Numerical Degrees to Ordered Labels, With Possible Applications to Dark Matter and Dark Energy

Olga Kosheleva, Vladik Kreinovich,
Martha Osegueda Escobar, Kimberly Kato
University of Texas at El Paso
500 W. University
El Paso, TX 79968, USA
olgak@utep.edu, vladik@utep.edu
mcoseguedaescobar@miners.utep.edu,
kekato@miners.utep.edu

Abstract—Experts often describe their estimates by using words from natural language, i.e., in effect, sorted labels. To efficiently represent the corresponding expert knowledge in a computer-based system, we need to translate these labels into a computer-understandable language, i.e., into numbers. There are many ways to translate labels into numbers. In this paper, we propose to select a translation which is the most robust, i.e., which preserves the order between the corresponding numbers under the largest possible deviations from the original translation. The resulting formulas are in good accordance with the translation coming from the Laplace’s principle of sufficient reason, and – somewhat surprisingly – with the current estimates of the proportion of dark matter and dark energy in our Universe.

I. FORMULATION OF THE PROBLEM

Ordered labels are ubiquitous. In many real-life situations ranging from medicine to geology to piloting planes, we rely on human experts. In particular, we rely on the expert’s ability to estimate the values of different relevant properties and to make decision based on these estimates.

In some cases, an expert can express his/her estimate in numerical terms: a distance to the car nearby is about 50 m, the probability of finding oil in this area is about 70%, etc. However, frequently, the expert can only describe his/her estimate by using words from natural language: the nearest car is *somewhat close*, it is *highly probable* that the area contains oil, etc.

The corresponding natural-language terms serve as *labels* making different values. Usually, the set of labels is *linearly ordered*: for every two labels, we know which one describes largest degree of distance or probability. For example, “far away” is farther than “somewhat far away”, “highly probable” means higher probability than “somewhat probable”, etc.

It is desirable to translate these labels into numbers. Some experts are better than others. It would be great if we could have the world’s best doctors treat all the patients, and the world’s

best pilots to fly all the planes. However, there are too many patients (and too many planes), so it is not possible to always use the best experts.

It is therefore desirable to design computer-based systems that would incorporate the knowledge of the best experts – and thus, provide relevant advice in situations when the top experts are not available. Such systems are known as *expert systems*.

Computers, however, have been originally designed to process numbers, not words from natural language. Modern computers are still much better in processing numbers than in processing words from natural language. Thus, to make computer-based systems efficient, it is desirable to translate labels into numbers.

Comment. The need for such a translation is one of the main ideas behind *fuzzy logic* (see, e.g., [5], [11], [12]), a methodology that has been specifically designed to translate expert (and commonsense) knowledge into numbers and to utilize the resulting formalized knowledge.

Resulting problem. We have a finite number of labels ℓ_1, \dots, ℓ_n ; these labels are sorted:

$$\ell_1 < \dots < \ell_n.$$

We would like to assign, to each label ℓ_i , a number r_i so that these numbers preserve the order:

$$r_1 < \dots < r_n.$$

When we estimate probabilities, these numbers should be from the interval $[0, 1]$. When we estimate distances, we may have other intervals – but all these intervals can be transformed, by selecting an appropriate measuring unit, into an interval $[0, 1]$. Thus, without losing generality, we can assume that the values r_i are all from the interval $[0, 1]$.

The very fact that the expert uses an imprecise label instead of saying that the event is absolutely possible or absolutely impossible means that the corresponding labels are different from the absolute probabilities 0 and 1; thus,

$$r_0 \stackrel{\text{def}}{=} 0 < r_1 < \dots < r_n < r_{n+1} \stackrel{\text{def}}{=} 1. \quad (1)$$

The problem is that there are many different tuples r_i satisfying the inequality (1). Which one should we choose?

II. MAIN IDEA: ROBUSTNESS

Main idea. Since there is a lot of freedom in selecting the numbers r_i , there is no need to perform exact computations with these numbers – approximate computations will save us computation time, which is important in many time-critical practical situations. Thus, instead of the exact values $r_0, r_1, \dots, r_n, r_{n+1}$, we may end up with approximate values $r'_i \approx r_i$, which are accurate up to a certain accuracy $\varepsilon > 0$, i.e., for which

$$|r_i - r'_i| \leq \varepsilon \quad (2)$$

for all labels $i = 0, 1, \dots, n + 1$.

We can modify the numerical values r_i , but we need to make sure that the modified values still preserve the same order, i.e., that

$$r'_0 < r'_1 < \dots < r'_n < r'_{n+1}. \quad (3)$$

In mathematical terms, we thus want to make sure that the tuple r_i is *robust* with respect to such modifications.

Clearly, when the modifications are large enough, i.e., when the inaccuracy ε is large enough, then the order may be violated. For example, one can easily check that $|r - 0.5| \leq 0.5$ for all $r \in [0, 1]$. Thus, if we take $\varepsilon = 0.5$, then for the values $r'_0 = r'_1 = \dots = r'_n = r'_{n+1} = 0.5$ for which $|r_i - r'_i| \leq 0.5$, the order is not preserved.

The larger the value ε for which the order is guaranteed to be preserved, the less accurate computations are needed and thus, the faster are the resulting computations. From this viewpoint, it is reasonable to select a tuple r_i for which the corresponding value ε is the largest possible – i.e., the most robust tuple.

Towards a formalization of the main idea. Let us describe this requirement in precise terms.

Definition. Let $n \geq 1$ be a natural number.

- By an n -tuple, we mean a sequence of real numbers

$$r_0 < r_1 < \dots < r_n < r_{n+1}$$

from the interval $[0, 1]$.

- For each n -tuple r_i , its robustness degree $d(r)$ is the supremum of all the values $\varepsilon > 0$ for which, if $|r'_i - r_i| \leq \varepsilon$ for all i , then $r'_0 < r'_1 < \dots < r'_n < r'_{n+1}$.
- We say that an n -tuple r_i is the most robust if its robustness degree is the largest possible.

Proposition. The only most robust n -tuple is the tuple

$$r_i = \frac{i}{n+1},$$

for which

$$r_0 = 0 < r_1 = \frac{1}{n+1} < \dots < r_n = \frac{n}{n+1} < r_{n+1} = 1.$$

Proof.

1. Let us first prove that for each n -tuple $r = (r_1, \dots, r_n)$, its robustness degree $d(r)$ is equal to

$$d(r) = \min_i \frac{r_{i+1} - r_i}{2}.$$

To prove this statement, it is sufficient to prove two auxiliary statements:

- that if $\varepsilon < d(r)$, then every sequence r'_i for which

$$|r'_i - r_i| \leq \varepsilon$$

is sorted, and

- that if $\varepsilon \geq d(r)$, then there exists a sequence r'_i for which $|r'_i - r_i| \leq \varepsilon$ for all i but which is not sorted.

Let us prove these two statements one by one.

1.1. If $\varepsilon < d(r)$, then, from $|r'_i - r_i| \leq \varepsilon$ and $|r'_{i+1} - r_{i+1}| \leq \varepsilon$, we conclude that $r'_{i+1} \geq r_{i+1} - \varepsilon$ and $r'_i \leq r_i + \varepsilon$. Thus, we have

$$r'_{i+1} - r'_i \geq (r_{i+1} - \varepsilon) - (r_i + \varepsilon) = (r_{i+1} - r_i) - 2\varepsilon.$$

We assume that $\varepsilon < d(r) = \min_i \frac{r_{i+1} - r_i}{2}$, thus $\varepsilon < \frac{r_{i+1} - r_i}{2}$ and $2\varepsilon < r_{i+1} - r_i$, so $(r_{i+1} - r_i) - 2\varepsilon > 0$. Hence,

$$r'_{i+1} - r'_i \geq (r_{i+1} - r_i) - 2\varepsilon > 0,$$

i.e., indeed, $r'_{i+1} > r'_i$.

1.2. Let us now prove that if $\varepsilon \geq d(r)$, then there exists a sequence r'_i for which $|r'_i - r_i| \leq \varepsilon$ and which is no longer sorted. Indeed, let i_0 be an index for which the value of the difference $r_{i_0+1} - r_{i_0}$ is the smallest:

$$r_{i_0+1} - r_{i_0} = \min_i (r_{i+1} - r_i).$$

Let us take a sequence r'_i for which

$$r'_{i_0+1} = r'_{i_0} = \frac{r_{i_0} + r_{i_0+1}}{2}$$

and $r'_i = r_i$ for all other i . In this case,

$$|r'_{i_0+1} - r'_{i_0}| = |r'_{i_0} - r_{i_0}| = \frac{r_{i_0+1} - r_{i_0}}{2} =$$

$$\min_i \frac{r_{i+1} - r_i}{2} = d(r).$$

Since we assumed that $\varepsilon \geq d(r)$, we thus conclude that

$$|r'_{i_0+1} - r'_{i_0}| = |r'_{i_0} - r_{i_0}| = d(r) \leq \varepsilon.$$

On the other hand, here $r'_{i_0+1} = r'_{i_0}$, so the order is clearly not preserved.

The statement is proven.

2. For the sequence $r_i = \frac{i}{n+1}$, the above robustness degree is equal to

$$d(r) = \frac{1}{2(n+1)}.$$

Let us prove that:

- no other sequence has a larger robustness degree, and
- the above sequence is the only one with this robustness degree.

2.1. Let us first prove, by contradiction, that a larger robustness degree is impossible. Indeed, let us assume that there exists an n -tuple s_i for which $d(s) > \frac{1}{2(n+1)}$. By the above formula for the robustness degree, this means that

$$\min_i \frac{s_{i+1} - s_i}{2} > \frac{1}{2(n+1)},$$

thus

$$\frac{s_{i+1} - s_i}{2} > \frac{1}{2(n+1)}$$

for all i and therefore,

$$s_{i+1} - s_i > \frac{1}{n+1}. \quad (4)$$

Here, $s_0 \geq 0$ and $s_{n+1} \leq 1$, hence

$$1 \geq s_{n+1} - s_0 = (s_{n+1} - s_n) + \dots + (s_1 - s_0).$$

Due to the above inequality (4), we get

$$1 > \frac{1}{n+1} + \dots + \frac{1}{n+1} (n+1 \text{ times}) = 1,$$

i.e., $1 > 1$, a clear contradiction. The statement is proven.

2.2. Let us now prove that the sequence $r_i = \frac{i}{n+1}$ is the only one with the robustness degree $d(r) = \frac{1}{2(n+1)}$. Namely, we assume that an n -tuple r_i has this robustness degree, and we will prove that $r_i = \frac{i}{n+1}$ for all i .

To prove this equality, we will prove the following two inequalities: that $r_i \geq \frac{i}{n+1}$ and $r_i \leq \frac{i}{n+1}$. Let us prove these two inequalities one by one.

2.2.1. From the condition that $d(r) = \min_i \frac{r_{i+1} - r_i}{2}$, it follows that $\frac{r_{i+1} - r_i}{2} \geq d(r)$ for all i and thus, $r_{i+1} - r_i \geq 2d(r)$. In our case, this means that $r_{i+1} - r_i \geq \frac{1}{n+1}$.

Now, for every i , since $r_0 \geq 0$, we have

$$r_i \geq r_i - r_0 = (r_i - r_{i-1}) + \dots + (r_1 - r_0) \geq \frac{1}{n+1} + \dots + \frac{1}{n+1} (i \text{ times}) = \frac{i}{n+1}.$$

The first inequality is proven.

2.2.2. Similarly, since $r_{n+1} \leq 1$, we have

$$1 - r_i \geq r_{n+1} - r_i = (r_{n+1} - r_n) + \dots + (r_{i+1} - r_i) \geq$$

$$\frac{1}{n+1} + \dots + \frac{1}{n+1} (n+1-i \text{ times}) = \frac{n+1-i}{n+1} = 1 - \frac{i}{n+1}.$$

From

$$1 - r_i \geq 1 - \frac{i}{n+1},$$

we conclude that $r_i \geq \frac{i}{n+1}$. Thus, the second inequality is also proven, and so is the proposition.

III. RELATION TO LAPLACE'S PRINCIPLE OF SUFFICIENT REASON

What is this principle. Laplace's principle of sufficient reason states that if we have no reason to assume that one of the events is more probable than others, then it makes sense to assume that these events are equally probable.

How can this principle be applied to our situation. In our situation, this means that we consider all the tuples r_i for which $0 < r_1 < \dots < r_n < 1$ to be equally probable. It is therefore reasonable to select an *average* (mean) value of such a tuple.

The results are in good accordance with the robustness approach. It turns out that this approach leads exactly to the same tuple $r_i = \frac{i}{n+1}$ [2], [3], [4], [6], [7], [8], [9], [10], which makes us more confident that the above robustness approach makes sense.

However, robustness approach is easier to describe and analyze: no probabilities, no averages over all possible tuples – i.e., no integration, etc.

IV. POSSIBLE APPLICATION TO DARK MATTER AND DARK ENERGY

Dark matter and dark energy: a brief reminder. Only one force affects the motion of celestial objects: the gravity force. By observing trajectories of stars in galaxies, we can estimate the corresponding gravity force and thus, find the masses causing this force. On the other hand, when we add up the masses of all the observed celestial bodies within a galaxy, we get a much smaller number. The missing – invisible – mass is known as *dark matter*.

We can perform a similar analysis on the cosmological level, by comparing the cosmological dynamics with the overall mass of usual matter and dark matter. Again, it turns out that some mass is missing. The missing mass is called *dark energy*; see, e.g., [1].

Dark energy and dark matter: empirical data. According to the empirical data, dark energy constitutes approximately 68% of the Universe's mass, the remaining 32% is dark matter and usual matter.

Dark matter is divided into three different categories based on the velocity of the corresponding particles: cold, warm, and hot. In the past, hot dark matter was considered to be a prevalent form of it, but now, it turned out that its role is small (if at all). So, in the first approximation, we can ignore hot dark matter. The idea is that while the Universe started in a hot state, by now, after billions of years, it has cooled down.

So, similarly to the usual matter, most dark matter is cold, second is amount is warm dark matter. In this approximation, in addition to dark energy, we have three different types of matter: cold dark matter, warm dark matter, and the usual (baryonic) matter. The usual matter makes up between 4-5% of the overall mass of the Universe, it is smaller than the amounts of cold or warm dark matter.

How can we explain these numbers? Let us first consider the division of the Universe’s masses into dark energy and matter proper.

All we know is that there is more dark energy than dark matter. Here, we have $n = 2$ labels, two expert estimates. The above argument suggests that the larger label $r_2 > r_1$ corresponds to

$$r_2 = \frac{2}{2+1} = \frac{2}{3} \approx 67\%,$$

while the lower label $r_1 < r_2$ corresponds to

$$r_1 = \frac{1}{2+1} = \frac{1}{3} \approx 33\%.$$

This is amazingly close – with 1% accuracy – to the current estimates 68% and 32%.

Out of the 32% of matter, we have three different types: the usual matter, the cold dark matter, and the warm dark matter (listed in the increasing order). According to our description, the proportion of usual matter is thus proportional to $\frac{1}{4}$, of cold dark matter to $\frac{2}{4}$, and of warm dark matter to $\frac{3}{4}$. Thus, the relative proportion of the usual matter is equal to the ratio

$$\frac{\frac{1}{4}}{\frac{1}{4} + \frac{2}{4} + \frac{3}{4}} = \frac{1}{1+2+3} = \frac{1}{6}.$$

One sixth of 32% is approximately 5%, which is in line with the above empirical estimate for the amount of usual matter.

Conclusion. Thus, the above robustness-based analysis is in good accordance with the observed amounts of dark matter and dark energy.

ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation grants HRD-0734825 and HRD-1242122 (Cyber-ShARE Center of Excellence) and DUE-0926721, and by an award “UTEP and Prudential Actuarial Science Academy and Pipeline Initiative” from Prudential Foundation.

REFERENCES

- [1] P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, et al., “Planck 2013 results. I. Overview of products and scientific results”, *Astronomy and Astrophysics*, 2014, Vol. 571, paper A1.
- [2] M. Ahsanullah, V. B. Nevzorov, and M. Shakil, *An Introduction to Order Statistics*, Atlantis Press, Paris, 2013.
- [3] B. C. Arnold, N. Balakrishnan, and H. N. Nagaraja, *A First Course in Order Statistics*, Society of Industrial and Applied Mathematics (SIAM), Philadelphia, Pennsylvania, 2008.
- [4] H. A. David and H. N. Nagaraja, *Order Statistics*, Wiley, New York, 2003.
- [5] G. Klir and B. Yuan, “Fuzzy Sets and Fuzzy Logic”, Prentice Hall, Upper Saddle River, New Jersey, 1995.
- [6] O. Kosheleva, V. Kreinovich, J. Lorkowski, and M. Osegueda, “How to Transform Partial Order Between Degrees into Numerical Values”, *Proceedings of the 2016 IEEE International Conferences on Systems, Man, and Cybernetics SMC’2016*, Budapest, Hungary, October 9–12, 2016.
- [7] J. Lorkowski and V. Kreinovich, “Interval and Symmetry Approaches to Uncertainty – Pioneered by Wiener – Help Explain Seemingly Irrational Human Behavior: A Case Study”, *Proceedings of the 2014 Annual Conference of the North American Fuzzy Information Processing Society NAFIPS’2014*, Boston, Massachusetts, June 24–26, 2014.
- [8] J. Lorkowski and V. Kreinovich, “Likert-type fuzzy uncertainty from a traditional decision making viewpoint: how symmetry helps explain human decision making (including seemingly irrational behavior)”, *Applied and Computational Mathematics*, 2014, Vol. 13, No. 3, pp. 275–298.
- [9] J. Lorkowski and V. Kreinovich, “Granularity Helps Explain Seemingly Irrational Features of Human Decision Making”, In: W. Pedrycz and S.-M. Chen (eds.), *Granular Computing and Decision-Making: Interactive and Iterative Approaches*, Springer Verlag, Cham, Switzerland, 2015, pp. 1–31.
- [10] J. Lorkowski and V. Kreinovich, “Fuzzy Logic Ideas Can Help in Explaining Kahneman and Tversky’s Empirical Decision Weights”, In: L. Zadeh et al. (Eds.), *Recent Developments and New Direction in Soft-Computing Foundations and Applications*, Springer Verlag, to appear.
- [11] H. T. Nguyen and E. A. Walker, *A First Course in Fuzzy Logic*, Chapman and Hall/CRC, Boca Raton, Florida, 2006.
- [12] L. A. Zadeh, “Fuzzy sets”, *Information and Control*, 1965, Vol. 8, pp. 338–353.