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, and Kimberly Kato
University of Texas at El Paso
El Paso, TX 79968, USA
olgak@utep.edu, vladik@utep.edu
mcoseguedaescobar@miners.utep.edu, kekato@miners.utep.edu

1. 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 quantities, 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 langauge.
- *Example:* the nearest car is *somewhat close*.

2. Need for Expert Systems

- The corresponding natural-language terms serve as *labels* making different values.
- Usually, the set of labels is *linearly ordered*:
  - "far away" is farther than "somewhat far away",
  - "highly probable" means higher probability than "somewhat probable"', etc.
- Some experts are better than others; it would 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 planes), so it is not possible to always use the best experts.

3. Need to Translate Labels into Numbers

- It is desirable to design computer-based systems incorporating the knowledge of the best experts.
- 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 natural language.
- Thus, to make computer-based systems efficient, it is desirable to translate labels into numbers.
- The need for such a translation is one of the main ideas behind *fuzzy logic*.

4. Resulting Problem

- We have a finite number of labels $\ell_1, \ldots, \ell_n$.
- These labels are sorted: $\ell_1 < \ldots < \ell_n$.
- We would like to assign, to each label ℓ_i, a number $r_i \in [0, 1]$ so that these numbers preserve the order:

$$r_0 \stackrel{\text{def}}{=} 0 < r_1 < \ldots < r_n < r_{n+1} \stackrel{\text{def}}{=} 1.$$

- The problem is that there are many different tuples r_i satisfying this inequality.
- Which one should we choose?

5. Main Idea: Robustness

- There is a lot of freedom in selecting the numbers r_i.
- So, 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_i, we may end up with approximate values $r'_i \approx r_i$: $|r_i - r'_i| \le \varepsilon$.
- We need to make sure that the modified values preserve the same order: $r'_0 < r'_1 < \ldots < r'_n < r'_{n+1}$.
- In mathematical terms, we thus want to make sure that the tuple r_i is *robust* with respect to such modifications.

6. Main Idea (cont-d)

- When the inaccuracy ε is large enough, then the order may be violated.
- The larger the value ε for which the order is guaranteed to be preserved:
  - the less accurate computations are needed,
  - thus, the faster are the resulting computations.
- So, it is reasonable to select a tuple r_i for which the corresponding value ε is the largest possible.
- By an *n-tuple*, we mean a sequence of real numbers $r_0 < r_1 < \ldots < r_n < r_{n+1}$ 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 $|r_i' - r_i| \le \varepsilon$ for all i, then $r_0' < r_1' < \ldots < r_n' < r_{n+1}'$.

7. Main Result

- 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 $r_i = \dfrac{i}{n+1}$, for which*

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

- **Proof.** Let us first prove that for each n-tuple $r = (r_1, \ldots, r_n)$, $d(r) = \min_i \dfrac{r_{i+1} - r_i}{2}$.
- To prove this statement, it is sufficient to prove two auxiliary statement:
    - that if $\varepsilon < d(r)$, then every sequence r'_i for which $|r'_i - r_i| \le \varepsilon$ is sorted, and
    - that if $\varepsilon \ge d(r)$, then there exists a sequence r'_i for which $|r'_i - r_i| \le \varepsilon$ for all i but which is not sorted.

8. Proof (cont-d)

- Let us prove these two statements one by one.
- If $\varepsilon < d(r)$, then, from $|r'_i - r_i| \le \varepsilon$ and $|r'_{i+1} - r_{i+1}| \le \varepsilon$, we conclude that $r'_{i+1} \ge r_{i+1} - \varepsilon$ and $r'_i \le r_i + \varepsilon$.
- So, $r'_{i+1} - r'_i \ge (r_{i+1} - \varepsilon) - (r_i + \varepsilon) = (r_{i+1} - r_i) - 2\varepsilon$.
- We assume that $\varepsilon < d(r) = \min\limits_i \dfrac{r_{i+1} - r_i}{2}$, thus $\varepsilon < \dfrac{r_{i+1} - r_i}{2}$ and $2\varepsilon < r_{i+1} - r_i$, so $(r_{i+1} - r_i) - 2\varepsilon > 0$.
- Hence, $r'_{i+1} - r'_i \ge (r_{i+1} - r_i) - 2\varepsilon > 0$, i.e., indeed, $r'_{i+1} > r'_i$.
- Let us now prove that if $\varepsilon \ge d(r)$, then there exists an unsorted r'_i for which $|r'_i - r_i| \le \varepsilon$.
- Indeed, let i_0 be an index for which the difference $r_{i+1} - r_i$ is the smallest: $r_{i_0+1} - r_{i_0} = \min\limits_i (r_{i+1} - r_i)$.

9. Proof (cont-d)

- Let us take a sequence r'_i for which $r'_{i_0+1} = r'_{i_0} = \dfrac{r_{i_0} + r_{i_0+1}}{2}$ and $r'_i = r_i$ for all other i.
- Then, $|r'_{i_0+1} - r_{i_0+1}| = |r'_{i_0} - r_{i_0}| = \dfrac{r_{i_0+1} - r_{i_0}}{2} =$

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

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

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

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

10. Proof: Part 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 sequence has a larger robustness degree, and
    - the above sequence is the only one with this robustness degree.
- 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 $d(s) = \min\limits_i \frac{s_{i+1} - s_i}{2} > \frac{1}{2(n+1)}$.
- Thus $s_{i+1} - s_i > \frac{1}{n+1}$ for all i.

11. Proof: Part 2 (cont-d)

- Here, $s_0 \ge 0$ and $s_{n+1} \le 1$, hence

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

- Due to the above inequality, we get

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

  i.e., $1 > 1$, a clear contradiction.

- The statement is proven.

- Let us now prove that the sequence $r_i = \dfrac{i}{n+1}$ is the only one with the robustness degree $d(r) = \dfrac{1}{2(n+1)}$.

- We assume that an n-tuple r_i has this robustness degree, and we will prove that $r_i = \dfrac{i}{n+1}$ for all i.

12. Proof: Part 2 (cont-d)

- To prove this equality, we will prove the following two inequalities: that $r_i \ge \dfrac{i}{n+1}$ and $r_i \le \dfrac{i}{n+1}$.
- Let us prove these two inequalities one by one.
- From $d(r) = \min\limits_i \dfrac{r_{i+1} - r_i}{2}$, it follows that $\dfrac{r_{i+1} - r_i}{2} \ge d(r)$ for all i and thus, $r_{i+1} - r_i \ge 2d(r)$.
- In our case, this means that $r_{i+1} - r_i \ge \dfrac{1}{n+1}$.
- Now, for every i, since $r_0 \ge 0$, we have

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

- The first inequality is proven.

Ordered Labels are . . .
Need for Expert Systems
Need to Translate . . .
Resulting Problem
Main Idea: Robustness
Main Result
Relation to Laplace's . . .
Dark Matter and Dark . . .
How Can We Explain . . .
Home Page
Title Page
◀◀ ▶▶
◀ ▶
Page 13 of 22
Go Back
Full Screen
Close
Quit

13. Proof: Part 2 (final)

- Similarly, since $r_{n+1} \le 1$, we have

$$1 - r_i \ge r_{n+1} - r_i = (r_{n+1} - r_n) + \ldots + (r_{i+1} - r_i) \ge$$

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

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

14. Relation to Laplace's Principle of Sufficient Reason

- 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.
- In our situation, we consider all the tuples r_i for which $0 < r_1 < \ldots < r_n < 1$ to be equally probable.
- It is therefore reasonable to select an *average* (mean) value of such a tuple.
- It turns out that this approach leads exactly to the same tuple $r_i = \dfrac{i}{n+1}$.
- This makes us more confident that the above robustness approach makes sense.

15. Dark Matter and Dark Energy

- 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 similarly compare the cosmological dynamics with the overall mass of usual matter and dark matter.

16. Dark Matter and Dark Energy (cont-d)

- Again, it turns out that some mass is missing. The missing mass is called *dark energy*.
- 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.
- Now, it turned out that its role is small (if at all).

17. Dark Matter and Dark Energy (cont-d)

- The idea is that while the Universe started in a hot state, by now, 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.

18. 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
  $$r_2 = \frac{2}{2+1} = \frac{2}{3} \approx 67\%,$$
  - 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%.

19. Explanation (cont-d)

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

20. Explanation (final)

- 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 $\approx$ 5%, in line with the above empirical estimate for the amount of usual matter.
- So, robustness-based analysis is in accordance with the observed amounts of dark matter and dark energy.

21. 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 from Prudential Foundation.