IV-я Международная научная конференция ФКН
«Математическое и компьютерное моделирование»,
Россия, Омск, ОмГУ, 11 ноября 2016 г. – С. 1–5. УДК XXX

Kimberly Kato
Department of Computer Science
University of Texas at El Paso, El Paso, Texas 79968, USA
kekato@miners.utep.edu

HOW TO ASSIGN NUMERICAL VALUES TO PARTIALLY ORDERED LEVELS OF CONFIDENCE: ROBUSTNESS APPROACH

Outline. In many practical situations, expert's levels of confidence are described by words from natural language, and these words are only partially ordered. Since computers are much more efficient processing numbers than words, it is desirable to assign numerical values to these degrees. Of course, there are many possible assignments that preserve order between words. It is reasonable to select an assignment which is the most robust, i.e., for which the largest possible deviation from the numerical values still preserves the order. In this paper, we describe such assignments for situations when we have 2, 3, and 4 different words.

Need to assign numerical values to levels of confidence. In many cases, it is desirable to describe experts' knowledge in a computer-understandable form. Experts are often not 100% confident in their statements; the corresponding degrees of confidence are an important part of their knowledge. It is therefore desirable to describe these levels of confidence in a computer-understandable form.

Experts often describe their levels of confidence by using words from a natural language, such as "most probably", "usually", etc. Since computers are much more efficient when they process numbers than when they process words, it is desirable to describe these levels of confidence by numbers. In other words, it is desirable to assign numerical values to different levels of certainty.

These numerical values are usually selected from the interval $[0, 1]$, so that 1 corresponds to complete certainty, and 0 to full certainty that the statement is true.

There is an order $\prec$ between levels, with $a \prec b$ meaning that level b corresponds to higher confidence than level a. This order is often partial, i.e., there exist levels a and b for which it is not clear which of them corresponds to higher confidence. It is reasonable to assign degree in such a way that if $a \prec b$, then the degree assigned to level b is larger than the degree assigned to level a. Also, all assigned degrees should be strictly between 0 and 1 – since they describe different levels of certainty, not absolute certainty.

Notations. For simplicity, let us number all levels by 1, 2, ..., n. To these levels, we add ideal levels 0 (absolutely false) and $n+1$ (absolutely true), for which $o \prec i \prec n+1$ for all i from 1 to n. Let us denote the numerical value assigned to the i-th level by $n_i \in [0,1]$.

In these terms, our requirement means that $i \prec j$ implies $n_i < n_j$.

How to assign? There are many way to assign numbers to levels. For example, if we have $n = 2$ levels with $1 \prec 2$, then possible assignments are possible tuples (n_0, n_1, n_2, n_3) for which $n_0 < n_1 < n_2 < n_3$. Of course, there are many such tuples. Which of the possible assignments should we select?

Robustness as a possible criterion. Computers are approximate machines. The higher accuracy we need, the more digits we should keep in our computations, and thus, the slower are these computations. Therefore, to speed up computations, we would like to store as few digits as possible, i.e., to replace the original values with approximate ones.

We want to make sure that this approximation preserves the order, i.e., that if we replace the original values n_i with approximate values n'_i for which $|n'_i - n_i| \le \varepsilon$ for the corresponding accuracy ε, we will still have the same order between the new values n'_i as between the old values. So, we want the numerical assignment which is, in this sense, *robust*.

The larger ε, the fewer digits we can keep and thus, the faster the computations. Thus, it is desirable to select the assignment for which the robustness ε is the largest possible. In precise terms, we want to select numbers n_i for which $i \prec j$ implies $n'_i < n'_j$ whenever $|n'_i - n_i| \le \varepsilon$ and $|n'_j - n_j| \le \varepsilon$ – for the largest possible value ε.

If we have two arrangements with the same ε, but one of them allows for larger deviations of at least one of the values n_i than the other one, then we should select this one – since it is more robust.

What is known: case of a linear order. In [1], we have shown that for the case of linear order, when $1 \prec 2 \prec \ldots \prec n$, the most robust assignment is $n_i = \dfrac{i}{n+1}$, with the robustness $\varepsilon = \dfrac{1}{2(n+1)}$.

What we do in this paper. In this paper, we extend this result to partially ordered sets with up to 4 elements.

1-element set. This case is the easiest, since a 1-element set is, by definition, linearly ordered. So, in this case, we assign $n_1 = 1/2$.

2-element set.

- If the two elements are ordered ($1 \prec 2$), then we assign $n_1 = 1/3$ and $n_2 = 2/3$.
- If the elements are not related, then the most robust assignment is when $n_1 = n_2 = 1/2$.

3-element set. In this case, let us analyze different possible cases based on the number of *minimal* elements, i.e., elements which are not preceded by any others.

Case of 3 minimal elements. In this case, the three elements 1, 2, and 3 are unrelated, so the most robust assignment is $n_1 = n_2 = n_3 = 1/2$, with degree of robustness $1/4$.

Case of 2 minimal elements. Without losing generality, let us assume that 1 and 2 are minimal elements. Since the element 3 is not minimal, it has to have preceding elements. There are two subcases here: when both elements 1 and 2 are preceding and when only one of them is preceding; in the second case, without losing generality, we can assume that $1 \prec 3$.

- If $1 \prec 3$ and $2 \prec 3$, then we should take $n_1 = n_2 = 1/3$ and $n_3 = 2/3$.
- If $1 \prec 3$ and 2 is not related, we should take $n_1 = 1/3$, $n_2 = 2/3$, and $n_2 = 1/2$.

Comment. We get the same robustness level $\varepsilon = 1/6$ for all possible values $n_2 \in [1/3, 2/3]$. We select $n_2 = 1/2$ since for this value, the largest possible deviation of n_2 preserves the order when all other values are ε-disturbed.

Case of a single minimal element. Without losing generality, we can assume that this minimal element is 1. Since 2 and 3 are not minimal, they have to have a preceding element, and since the only minimal element is 1, they have to have 1 as preceding. There are subcases: when 2 and 3 are unrelated and when they are related; in the second subcase, without losing generality, we can assume that $2 \prec 3$.

- If $1 \prec 2$ and $1 \prec 3$, then we should have $n_1 = 1/3$ and

$$n_2 = n_3 = 2/3.$$

- If $1 \prec 2 \prec 3$, then we have a linear order, so we should have $n_1 = 1/4$, $n_2/1/2$, and $n_3 = 3/4$.

4 elements, all 4 minimal. In this case, $n_1 = n_2 = n_3 = n_4 = 1/2$.

4 elements, 3 minimal. In this case, we have subcases depending on how many minimal elements precede the fourth (non-minimal) one.

- If we only have $1 \prec 4$, then $n_1 = 1/3$, $n_2 = n_3 = 1/2$, and $n_4 = 2/3$.
- If $1 \prec 4$ and $2 \prec 4$, then $n_1 = n_2 = 1/3$, $n_3 = 1/2$, and $n_4 = 2/3$.
- If $1 \prec 4$, $2 \prec 4$, and $3 \prec 4$, then $n_1 = n_2 = n_3 = 1/3$ and $n_4 = 2/3$.

4 elements, 2 minimal. Here, we have subcases depending on whether non-minimal elements 3 and 4 are related, and whether both minimal element precede something.

- If 3 and 4 are unrelated and both 1 and 2 precede others, then $n_1 = n_2 = 1/3$ and $n_3 = n_4 = 2/3$.
- If 3 and 4 are unrelated but only one 1 and 2 precedes others (e.g., 1), then $n_1 = 1/3$, $n_2 = 1/2$, and $n_3 = n_4 = 2/3$.
- If 3 and 4 are related, then, without losing generality, we can assume that $3 \prec 4$. If both 1 and 2 precede others, then $n_1 = n_2 = 1/4$, $n_3 = 1/2$, and $n_4 = 3/4$.
- If $3 \prec 4$ and 2 does not precede anything, then $n_1 = 1/4$, $n_2 = n_3 = 1/2$, and $n_4 = 3/4$.

4 elements, 1 minimal. The minimal element 1 should precede all other elements, and for other elements, we have the same possibilities as for the 3-element configuration. So, we get the following results.

- If $1 \prec 2$, $1 \prec 3$, and $1 \prec 4$, then $n_1 = 1/3$ and $n_2 = n_3 = n_4 = 2/3$.
- If $1 \prec 2 \prec 4$ and $1 \prec 3 \prec 4$, then $n_1 = 1/4$, $n_2 = n_3 = 1/2$, and $n_4 = 3/4$.

- If $1 \prec 2 \prec 4$ and $1 \prec 3$, then the most robust assignment is $n_1 = 1/4$, $n_2 = 1/2$, $n_3 = 5/8$, and $n_4 = 3/4$. (Here, n_3 is in the midpoint between n_1 and $n_5 = 1$, to guarantee maximal robustness with respect to changing n_3.)
- If $1 \prec 2 \prec 3$ and $1 \prec 2 \prec 4$, then $n_1 = 1/4$, $n_2 = 1/2$, and $n_3 = n_4 = 3/4$.
- Finally, if $1 \prec 2 \prec 3 \prec 4$, then $n_1 = 1/5$, $n_2 = 2/5$, $n_3 = 3/5$, and $n_4 = 4/5$.

Литература.

1. *Kosheleva, O., Kreinovich, V., Osegueda Escobar, M., and Kato, K.* Towards the most robust way of assigning numerical degrees to ordered labels, with possible applications to dark matter and dark energy // Proceedings of the 2016 Annual Conference of the North American Fuzzy Information Processing Society NAFIPS'2016, El Paso, Texas, October 31 – November 4, 2016, to appear.