3-2015

Coming Up with a Good Question Is Not Easy: A Proof

Joe Lorkowski
University of Texas at El Paso, lorkowski@acm.org

Luc Longpre
University of Texas at El Paso, longpre@utep.edu

Olga Kosheleva
University of Texas at El Paso, olgak@utep.edu

Salem Benferhat
Université d'Artois, benferhat@cril.univ-artois.fr

Recommended Citation

Lorkowski, Joe; Longpre, Luc; Kosheleva, Olga; and Benferhat, Salem, "Coming Up with a Good Question Is Not Easy: A Proof" (2015). *Departmental Technical Reports (CS).* Paper 901.
http://digitalcommons.utep.edu/cs_techrep/901

Coming Up with a Good Question Is Not Easy: A Proof

Joe Lorkowski, Luc Longpré, Olga Kosheleva
University of Texas at El Paso
500 W. University
El Paso, TX 79968, USA
lorkowski@computer.org, longpre@utep.edu, olgak@utep.edu

Salem Benferhat
Centre de Recherche en Informatique de Lens CRIL
Université d'Artois, F62307 Lens Cedex, France
benferhat@cril.univ-artois.fr

Abstract—**Ability to ask good questions is an important part of learning skills. Coming up with a good question, a question that can really improve one's understanding of the topic, is not easy. In this paper, we prove – on the example of probabilistic and fuzzy uncertainty – that the problem of selecting of a good question is indeed hard.**

I. FORMULATION OF THE PROBLEM

Asking good questions is important. Even after a very good lecture, some parts of the material remains not perfectly clear. A natural way to clarify these parts is to ask questions to the lecturer.

Ideally, we should be able to ask a question that immediately clarifies the desired part of the material. Coming up with such good questions is an important part of learning process, it is a skill that takes a long time to master.

Coming up with good questions is not easy: an empirical fact. Even for experienced people, it is not easy to come up with a good question, i.e., with a question that will maximally decrease uncertainty.

What we do in this paper. In this paper, we prove that the problem of designing a good question is indeed computationally difficult (NP-hard).

We will show this both for probabilistic and for fuzzy uncertainty. Specifically, we will prove NP-hardness for the simplest types of questions – for "yes"-"no" questions for which the answer is "yes" or "no". Since already designing such simple questions is NP-hard, any more general problem (allowing more complex problems) is NP-hard as well.

II. TOWARDS DESCRIBING THE PROBLEM IN PRECISE TERMS: GENERAL CASE

What is uncertainty: a general description. A complete knowledge about any area – be it a physical system or an algorithm – would mean that we have the full description of the corresponding objects. From this viewpoint, uncertainty means that several different variants are consistent with our (partial) knowledge, and we are not sure which of these variants is true.

In the following text, we will denote possible variants by $v_1, v_2, \ldots, v_n$, or, if this does not cause any ambiguity, simply by $1, 2, \ldots, n$.

How to describe a "yes"-"no" question in these terms. A "yes"-"no" question is a question an answer to which eliminates possible variants. In general, this means that after we get the answer to this question, instead of the original set $\{1, \ldots, n\}$ of possible variants, we have a smaller set:

- if the answer is "yes", then we are limited to the set $Y \subset \{1, \ldots, n\}$ of all the variants which are consistent with the "yes"-answer;

- if the answer is "no", then we are limited to the set $N \subset \{1, \ldots, n\}$ of all the variants which are consistent with the "no"-answer.

These two sets are complements to each other.

Examples. In some cases, we are almost certain about a certain variant, i.e., variant v_1. In this case, a natural question to ask if whether this understanding is correct. For this question:

- the "yes"-set Y consists of the single variant v_1, while

- the "no"-set $\{2, \ldots, n\}$ contains all other variants.

In other cases, we are completely unclear about the topic, e.g., we are completely unclear what is the numerical value of a certain quantity, we are not even sure whether this value is positive or non-positive. In this case, a natural question is to ask whether the actual value is positive. In such situation, each of the sets Y and N contain approximately a half of all original variants.

III. FORMULATION OF THE PROBLEM IN PRECISE TERMS: CASE OF PROBABILISTIC UNCERTAINTY

Probabilistic approach to describing uncertainty: a description. In the probabilistic approach, we assign a probability $p_i \geq 0$ to each of the possible variants, so that these probabilities add up to 1: $\sum_{i=1}^{n} p_i = 1$. The probability p_i, for example, may describe the frequency is which the i-th variant turned out to be true in similar previous situations.

How to quantify an amount of uncertainty: probabilistic case. In the case of probabilistic uncertainty, there is a well-established way to gauge the amount of uncertainty: namely, the entropy [9], [18]

$$S = -\sum_{i=1}^{n} p_i \cdot \ln(p_i). \qquad (1)$$

This is a good estimate for the amount that we want to decrease by asking an appropriate question.

How do we select a question: idea. We would thus like to find the question that maximally decreases the uncertainty. Since in the probabilistic case, uncertainty is measured by entropy, we thus want to find a question that maximally decreases entropy.

How the answer changes the entropy. Once we know the answer to our "yes"-"no" question, the probabilities change.

If the answer was "yes", this means that the variants from the "no"-set N are no longer possible. For such variants $i \in N$, the new probabilities are 0s: $p_i' = 0$. For variants from the "yes"-set Y, the new probability is the conditional probability under the condition that the variant is in the "yes"-set, i.e.,

$$p_i' = p(i \mid Y) = \frac{p_i}{p(Y)}, \qquad (2)$$

where the probability $p(Y)$ of the "yes"-answer is equal to the sum of the probabilities of all the variants that lead to the "yes"-answer:

$$p(Y) = \sum_{i \in Y} p_i. \qquad (3)$$

Based on these new probabilities, we can compute the new entropy value

$$S' = -\sum_{i \in Y} p_i' \cdot \ln(p_i'). \qquad (4)$$

On the other hand, if the answer was "no", this means that the variants from the "yes"-set Y are no longer possible. For such variants $i \in Y$, the new probabilities are 0s: $p_i'' = 0$. For variants from the "no"-set N, the new probability is the conditional probability under the condition that the variant is in the "yes"-set, i.e.,

$$p_i'' = p(i \mid N) = \frac{p_i}{p(N)}, \qquad (5)$$

where the probability $p(N)$ of the "no"-answer is equal to the sum of the probabilities of all the variants that lead to the "no"-answer:

$$p(N) = \sum_{i \in N} p_i. \qquad (6)$$

Based on these new probabilities, we can compute the new entropy value

$$S'' = -\sum_{i \in N} p_i'' \cdot \ln(p_i''). \qquad (7)$$

In the case of the "yes" answer, the entropy decreases by the amount $S - S'$. In the case of the "no"-answer, the entropy decreases by the amount $S - S''$. We know the probability $p(Y)$ of the "yes"-answer and we know the probability $p(N)$ of the "no"-answer. Thus, we can estimate the expected decrease in uncertainty as

$$\overline{S}(Y) = p(Y) \cdot (S - S') + p(N) \cdot (S - S''). \qquad (8)$$

Thus, we arrive at the following formulation of the problem in precise terms.

Formulation of the problem in precise terms.

- We are *given* the probabilities $p_1, \ldots, p_n$ for which

$$\sum_{i=1}^{n} p_i = 1.$$

- We need to *find* a set $Y \subset \{1, \ldots, n\}$ for which the expected decreased in uncertainty $\overline{S}(Y)$ is the largest possible.

Here, $\overline{S}(Y)$ is described by the formula (8), and the components of this formula are described in formulas (1)-(7).

IV. MAIN RESULT: PROBABILISTIC CASE

Formulation of the main result. Our main result is that the above problem – of coming up with the best possible question – is NP-hard.

What is NP-hard: a brief reminder. In many real-life problems, we are looking for a string (or for a sequence of a priori bounded numbers) that satisfies a certain property. For example, in the *subset sum* problem, we are given positive integers $s_1, \ldots, s_n$ representing the weights, and we need to divide these weights into two groups with exactly the same weight. In precise terms, we need to find a set $I \subseteq \{1, \ldots, n\}$ for which

$$\sum_{i \in I} s_i = \frac{1}{2} \cdot \left(\sum_{i=1}^{n} s_i \right).$$

The desired set I can be described as a sequence of n 0s and 1s, in which the i-th term is 1 if $i \in I$ and 0 if $i \notin I$.

In principle, we can solve each such problem by simply enumerating all possible strings, all possible combinations of numbers, etc. For example, in the above case, we can try all 2^n possible subsets of the set $\{1, \ldots, n\}$; this way, if there is a set I with the desired property, we will find it. The problem with this approach is that for large n, the corresponding number 2^n of computational steps becomes unreasonably large. For example, for $n = 300$, the resulting computation time exceeds lifetime of the Universe.

So, a natural question is: when can we solve such problems in feasible time, i.e., in time that does not exceed a polynomial of the size of the input? It is not know whether all exhaustive-search problems can be thus solved – this is the famous $P \overset{?}{=} NP$ problem. Most computer science researchers believe that some exhaustive-search problems cannot be feasibly solved – but in general, this remains an open problem.

What is known is that some problems are the hardest (NP-hard) in the sense that any exhaustive-search problems can be feasibly reduced to this problem. This means that, unless all exhaustive-search problems can be feasibly solved (which most computer scientists believe to be impossible), this particular problem cannot be feasibly solved.

The above subset sum problem has been proven to be NP-hard, as well as many other similar problems; see, e.g., [17].

How can we prove NP-hardness. As we have mentioned, a problem is NP-hard if every other exhaustive-search problem

$\mathcal{Q}$ can be reduced to it. So, if we know that a problem $\mathcal{P}_0$ is NP-hard, then every $\mathcal{Q}$ can be reduced to it. Thus, if $\mathcal{P}_0$ can be reduced to our problem $\mathcal{P}$, then, by transitivity, any problem $\mathcal{Q}$ can be reduced to $\mathcal{P}$, i.e., that $\mathcal{P}$ is indeed NP-hard.

Thus, to prove that a given problem is NP-hard, it is sufficient to reduce one known NP-hard problem $\mathcal{P}_0$ to this problem $\mathcal{P}$.

What we will do. To prove that the problem $\mathcal{P}$ of selecting a good question is NP-hard, we will follow the above idea. Namely, we will prove that the subset sum problem $\mathcal{P}_0$ (which is known to be NP-hard) can be reduced to $\mathcal{P}$.

Let us simplify the expression for $\overline{S}(Y)$. To build the desired reduction, let us simplify the expression (8). This expression uses the entropies S' and S''. So, to get the desired simplification, we will start with simplifying the expressions (4) and (7) for S' and S''.

Simplifying the expression for S'. Substituting the expression (2) for p_i' into the formula (4), we get

$$S' = -\sum_{i \in Y} \frac{p_i}{p(Y)} \cdot \ln\left(\frac{p_i}{p(Y)}\right).$$

All the terms in this sum are divided by $p(Y)$, so we can move this common denominator outside the sum:

$$S' = -\frac{1}{p(Y)} \cdot \left(\sum_{i \in Y} p_i \cdot \ln\left(\frac{p_i}{p(Y)}\right)\right).$$

The logarithm of the ratio is equal to the difference of logarithms, so we get

$$S' = -\frac{1}{p(Y)} \cdot \left(\sum_{i \in Y} p_i \cdot (\ln(p_i) - \ln(p(Y)))\right).$$

We can separate the terms proportional to $\ln(p_i)$ and to $\ln(p(Y))$ into two different sums. As a result, we get

$$S' = -\frac{1}{p(Y)} \cdot \left(\sum_{i \in Y} p_i \cdot \ln(p_i)\right) +$$

$$\frac{1}{p(Y)} \cdot \left(\sum_{i \in Y} p_i \cdot \ln(p(Y))\right). \tag{9}$$

In the second sum, the factor $\ln(p(Y))$ does not depend on i and can, thus, be moved out of the summation:

$$\sum_{i \in Y} p_i \cdot \ln(p(Y)) = \ln(p(Y)) \cdot \sum_{i \in Y} p_i.$$

Here, the sum $\sum_{i \in Y} p_i$ is simply equal to $p(Y)$, so

$$\sum_{i \in Y} p_i \cdot \ln(p(Y)) = \ln(p(Y)) \cdot p(Y). \tag{10}$$

Substituting the expression (10) into the formula (9), and cancelling the terms $p(Y)$ in the numerator and in the denominator, we conclude that

$$S' = -\frac{1}{p(Y)} \cdot \left(\sum_{i \in Y} p_i \cdot \ln(p_i)\right) + \ln(p(Y)). \tag{12}$$

Simplifying the expression for S''. Similarly, we get

$$S'' = -\frac{1}{p(N)} \cdot \left(\sum_{i \in N} p_i \cdot \ln(p_i)\right) + \ln(p(N)). \tag{13}$$

Resulting simplification of the expression for $\overline{S}(Y)$. Since $p(Y) + p(N) = 1$, the expression (8) for $\overline{S}(Y)$ can be alternatively described as

$$\overline{S}(Y) = S - (p(Y) \cdot S' + p(N) \cdot S''). \tag{14}$$

By using expressions (12) and (13) for S' and S'', we conclude that

$$p(Y) \cdot S' + p(N) \cdot S'' =$$

$$-\frac{p(Y)}{p(Y)} \cdot \left(\sum_{i \in Y} p_i \cdot \ln(p_i)\right) + p(Y) \cdot \ln(p(Y)) -$$

$$\frac{p(N)}{p(N)} \cdot \left(\sum_{i \in N} p_i \cdot \ln(p_i)\right) + p(N) \cdot \ln(p(N)) =$$

$$-\sum_{i \in Y} p_i \cdot \ln(p_i) - \sum_{i \in N} p_i \cdot \ln(p_i) +$$

$$p(Y) \cdot \ln(p(Y)) + p(N) \cdot \ln(p(N)). \tag{15}$$

Since Y and N are complements to each other, we have

$$-\sum_{i \in Y} p_i \cdot \ln(p_i) - \sum_{i \in N} p_i \cdot \ln(p_i) = -\sum_{i=1}^{n} p_i \cdot \ln(p_i) = S.$$

Thus, the formula (15) takes the form

$$p(Y) \cdot S' + p(N) \cdot S'' =$$

$$S + p(Y) \cdot \ln(p(Y)) + p(N) \cdot \ln(p(N)).$$

Therefore, the expression (14) takes the form

$$\overline{S}(Y) = -p(Y) \cdot \ln(p(Y)) - p(N) \cdot \ln(p(N)). \tag{16}$$

Here, $p(N) = 1 - p(Y)$, we have

$$\overline{S}(Y) = -p(Y) \cdot \ln(p(Y)) - (1 - p(Y)) \cdot \ln(1 - p(Y)). \tag{17}$$

Resulting reduction. We want to find the set Y that maximizes the expected decrease in uncertainty, i.e., that maximizes the expression (17). Thus, we need to select a question Y for which $p(Y) = 0.5$.

Let us show that a subset sum problem can be reduced to this problem. Indeed, let us assume that we are given n positive integers $s_1, \ldots, s_n$. Then, we can form n probabilities

$$p_i \stackrel{\text{def}}{=} \frac{s_i}{\sum_{j=1}^{n} s_j} \tag{18}$$

that add to 1. If for this problem, we can find a set Y for which $p(Y) = \sum_{i \in Y} p_i = 0.5$, then, due to the definition (18), for the original values s_i, we will have

$$\sum_{i \in Y} s_i = 0.5 \cdot \sum_{j=1}^{n} s_j. \tag{19}$$

This is exactly the solution to the subset sum problem. Vice versa, if we have a set Y for which the equality (19) is satisfied, then for the probabilities (18) we get $p(Y) = 0.5$.

Conclusion. The reduction shows that in the probabilistic case, the problem of coming up with a good question is indeed NP-hard.

V. FORMULATION OF THE PROBLEM IN PRECISE TERMS: CASE OF FUZZY UNCERTAINTY

Fuzzy approach to describing uncertainty: a description. In the fuzzy approach, we assign, to each variant i, its degree of possibility. The resulting fuzzy values are usually *normalized*, so that the largest of these values if equal to 1: $\max_i \mu_i = 1$; see, e.g., [11], [15], [19].

How to quantify amount of uncertainty: fuzzy case. In the case of fuzzy uncertainty, one of the most widely used ways to gauge uncertainty is to use an expression

$$S = \sum_{i=1}^{n} f(\mu_i), \tag{20}$$

for some strictly increasing continuous function $f(z)$ for which $f(0) = 0$; see, e.g., [16].

This is the amount that we want to decrease by asking an appropriate question.

How do we select a question: idea. We would thus like to find the question that maximally decreases the uncertainty. Since in the fuzzy case, uncertainty is measured by the expression (20), we thus want to find a question that maximally decreases the value of this expression.

How the answer changes the entropy. Once we know the answer to our "yes"-"no" question, the degrees of belief μ_i change.

If the answer was "yes", this means that the variants from the "no"-set N are no longer possible. For such variants $i \in N$, the new degrees are 0s: $\mu_i' = 0$. For variants from the "yes"-set Y, the new degree can be obtained by one of two different ways (see, e.g., [1], [2], [3], [4], [5], [6], [7], [8], [10]):

- in the *numerical approach*, we normalize the remaining degree so that the maximum is equal to 1, i.e., we take

$$\mu_i' = \frac{\mu_i}{\max_{j \in Y} \mu_j}; \tag{21}$$

- in the *ordinal approach*, we raise the largest values to 1, while keeping the other values unchanged:

$$\mu_i' = 1 \text{ if } \mu_i = \max_{j \in Y} \mu_j; \tag{22}$$

$$\mu_i' = \mu_i \text{ if } \mu_i < \max_{j \in Y} \mu_j. \tag{23}$$

Based on the new values μ_i', we compute the new complexity value

$$S' = \sum_{i \in Y} f(\mu_i'). \tag{24}$$

On the other hand, if the answer was "no", this means that the variants from the "yes"-set Y are no longer possible. For such variants $i \in Y$, the new degrees are 0s: $\mu_i'' = 0$. For variants from the "no"-set N, the new degree can be obtained by one of the same two different ways as in the case of the "yes" answer:

- in the *numerical approach*, we normalize the remaining degree so that the maximum is equal to 1, i.e., we take

$$\mu_i'' = \frac{\mu_i}{\max_{j \in N} \mu_j}; \tag{25}$$

- in the *ordinal approach*, we raise the largest values to 1, while keeping the other values unchanged:

$$\mu_i'' = 1 \text{ if } \mu_i = \max_{j \in N} \mu_j; \tag{26}$$

$$\mu_i'' = \mu_i \text{ if } \mu_i < \max_{j \in N} \mu_j. \tag{27}$$

Based on the new values μ_i', we compute the new complexity value

$$S'' = \sum_{i \in N} f(\mu_i''). \tag{28}$$

In the case of the "yes" answer, the uncertainty decreases by the amount $S - S'$. In the case of the "no"-answer, the uncertainty decreases by the amount $S - S''$. In this case, we do not know the probabilities of "yes" and '"no" answers, so we cannot estimate the expected decreased. What we can estimate is the *guaranteed* decrease

$$\overline{S}(Y) = \min(S - S', S - S''). \tag{29}$$

This value describes how much of a decrease we can guarantee if we use the "yes"-"no" answer corresponding to the set Y.

Thus, we arrive at the following formulation of the problem in precise terms.

Formulation of the problem in precise terms.

- We are *given* the degrees $\mu_1, \ldots, \mu_n$ for which

$$\max_i \mu_i = 1.$$

- We need to *find* a set $Y \subset \{1, \ldots, n\}$ for which the expected decreased in uncertainty $\overline{S}(Y)$ is the largest possible.

Here, $\overline{S}(Y)$ is described by the formula (29), and the components of this formula are described in formulas (20)-(28).

Comment. Strictly speaking, we need to solve two optimization problems:

- the problem corresponding to the numerical approach, and

- the problem corresponding to the ordinal approach.

VI.   Main Result: Fuzzy Case

Formulation of the main result. Our main result is that for both approaches (numerical and ordinal) the problem of coming up with the best possible question is NP-hard.

How we prove this result. Similarly to the probabilistic case, we prove this result by reducing the subset problem to this problem.

Reduction. Let $s_1, \ldots, s_m$ be positive integers. To solve the corresponding subset sum problem, let us select a small number $\varepsilon > 0$ and consider the following $n = m + 2$ degrees: $\mu_i = f^{-1}(\varepsilon \cdot s_i)$ for $i \leq m$ and $\mu_{m+1} = \mu_{m+2} = 1$, where $f^{-1}(z)$ denotes an inverse function to $f(z)$: $f^{-1}(z)$ is the value t for which $f(t) = z$.

For these values, we have three possible relations between the set Y and the variants $m + 1$ and $m + 2$:

- the first case is when the set Y contains both these variants;

- the second case is when the set Y contains none of these two variants, and

- the third case is when the set Y contains exactly one of these two variants.

Let us show that when ε is sufficiently small, then the largest guaranteed decrease is attained in the third case.

Indeed, one can easily check that

$$\overline{S}(Y) = \min(S - S', S - S'') = S - \max(S', S'');$$

Thus, the guaranteed decrease $\overline{S}(Y)$ is the largest when the maximum

$$\max(S', S'') \tag{30}$$

is the smallest.

In the first case, the values μ_i' contain two 1s, hence $S' = \sum_{i \in Y} f(\mu_i') \geq 2f(1)$. Thus, $\max(S', S'') \geq 2f(1)$.

In the second case, the values μ_i'' contain two 1s, hence $S'' = \sum_{i \in N} f(\mu_i'') \geq 2f(1)$. Thus, $\max(S', S'') \geq 2f(1)$.

In the third case, when one of the 1s in in Y and another one is in N, both sets Y and N contain 1s, so there is no need for normalization. Therefore, we have:

- $\mu_i' = \mu_i$ for $i \in Y$ and

- $\mu_i'' = \mu_i$ for $i \in N$.

Thus,

$$S' = \sum_{i \in Y} f(\mu_i) = f(1) + \sum_{i \in Y, i \leq m} f(\mu_i), \tag{31}$$

and

$$S'' = \sum_{i \in N} f(\mu_i) = f(1) + \sum_{i \in Y, i \leq m} f(\mu_i). \tag{32}$$

Due to our selection of μ_i, we have $f(\mu_i) = \varepsilon \cdot s_i$, so:

$$S' = \sum_{i \in Y} f(\mu_i) = f(1) + \sum_{i \in Y, i \leq m} \varepsilon \cdot s_i =$$

$$f(1) + \varepsilon \cdot \sum_{i \in Y, i \leq m} s_i, \tag{33}$$

and

$$S'' = \sum_{i \in N} f(\mu_i) = f(1) + \sum_{i \in N, i \leq m} \varepsilon \cdot s_i =$$

$$f(1) + \varepsilon \cdot \sum_{i \in N, i \leq m} s_i. \tag{34}$$

When $\varepsilon \cdot \sum_{i=1}^m s_i < f(1)$, we have $S' < 2f(1)$, $S'' < 2f(1)$, and therefore, $\max(S', S'') < 2f(1)$. Hence, for sufficiently small ε, the smallest possible value of the maximum (30) is indeed attained in the third case.

In this third case, due to (33) and (34), we have

$$\max(S', S'') =$$

$$f(1) + \varepsilon \cdot \max\left(\sum_{i \in Y, i \leq m} s_i, \sum_{i \in N, i \leq m} s_i \right).$$

The sets Y and N are complementary to each other, hence

$$\sum_{i \in Y, i \leq m} s_i + \sum_{i \in N, i \leq m} s_i = \sum_{i=1}^m s_i.$$

If the two sums $\sum_{i \in Y, i \leq m} s_i$ and $\sum_{i \in N, i \leq m} s_i$ are different, then one of them is larger that one half of the total sum $\sum_{i=1}^m s_i$; thus, the maximum $\max(S', S'')$ is also larger than one half of the total sum. The only way to get the smallest possible value – exactly one half of the total sum – is when the sums are equal to each other, i.e., when each sum is exactly one half of the total sum:

$$\sum_{i \in Y, i \leq m} s_i = \frac{1}{2} \cdot \left(\sum_{i=1}^m s_i \right).$$

This is exactly the solution to the subset problem. Thus, we have found the reduction of the known NP-hard subset sum problem to our problem of coming up with a good question - which implies that our problem is also NP-hard.

Conclusion. The reduction shows that in both fuzzy approaches, the problem of coming up with a good question is indeed NP-hard.

VII.   What Happens in the Interval-Valued Fuzzy Case

Need to consider interval-valued fuzzy sets. The usual $[0, 1]$-based fuzzy logic is based on the assumption that an expert can describe his or her degree of uncertainty by a number from the interval $[0, 1]$. In many practical situations, however, an expert is uncertain about his/her degree of uncertainty. In such situations, it is more reasonable to describe the expert's degree of certainty not by a single number, but by an interval which is a subinterval of the interval $[0, 1]$.

Such interval-valued fuzzy techniques have indeed led to many useful applications; see, e.g., [12], [13], [14].

The problem of selecting a good question is NP-hard under interval-valued fuzzy uncertainty as well. Indeed, the usual fuzzy logic is a particular case of interval-valued fuzzy logic – when all intervals are degenerate, i.e., are of the form $[a, a]$ for a real number a. It is easy to prove that if a particular case of a problem is NP-hard, the whole problem is also NP-hard.

Thus, since the problem of selecting a good question is NP-hard for the case of usual $[0, 1]$-based fuzzy uncertainty, it is also NP-hard for the more general case of interval-valued fuzzy uncertainty.

REFERENCES

[1] A. Ajroud and S. Benferhat, "An approximate algorithm for min-based possibilistic networks", *International Journal of Intelligent Systems*, to appear.

[2] R. Ayachi, N. Ben Amor, and S. Benferhat, "A generic framework for a compilation-based inference in probabilistic and possibilistic networks", *Information Sciences*, 2014, Vol. 257, pp. 342–356.

[3] R. Ayachi, N. Ben Amor, and S. Benferhat, "Inference using complied min-based possibilstic causal networks in the presence of interventions", *Fuzzy Sets and Systems*, 2014, Vol. 239, pp. 104–136.

[4] S. Benferhat, C. da Costa Pereira, and A. G. B. Tettamanzi, "Syntactic computation of hybrid possibilistic conditioning under uncertain inputs", In: F. Rossi, editor, *Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence IJCAI'2013*, Beijing, China, August 3–9, 2013.

[5] S. Benferhat, et al., "An intrusion detection and alert correlation approach based on revising probabilistic classifiers using expert knowledge", *Applied Intelligence*, 2012, Vol. 38, pp. 520–540.

[6] D. Dubois and H. Prade, "Possibility theory: qualitative and quantitative aspects", In: D. Gabbay and P. Smets, editors, *Handbook of Defeasible Reasoning and Uncertainty Management Systems. Vol. 1. Qnantified Rpresentation of Uncertainty and Imprecision*, Kluwer Academic Publ. Dordrecht, 1998, pp. 169–226.

[7] D. Dubois, J. Lang, and H. Prade, "Possibility theory: qualitative and quantitative aspects", In: D. Gabbay, C. J. Hogger, and J. A. Robinson, editors, *Handbook of Logic in Artifical Intelligence and Logic Programming. Vol. 3. Nonomonotonic Reasoning and Uncertain Reasoning*, Oxford University Press, New York, 1994, pp. 439–513.

[8] D. Dubois, S. Moral, and H. Prade, "Belief change rules in ordinal and numerical uncertainty theories", In: D. Gabbay and P. Smets, editors, *Handbook of Defeasible Reasioning and Uncertainty Management Systems. Vol. 1. Qnantified Rpresentation of Uncertainty and Imprecision*, Kluwer Academic Publ. Dordrecht, 1998, pp. 311–392.

[9] R. M. Gray, *Entropy and Information Theory*, Springer, Berlin, Heidelberg, New York, 2011.

[10] L. Gutierrez, M. Ceberio, V. Kreinovich, R. L. Gruver, M. Peña, M. J. Rister, A. Saldaña, J. Vasquez, J. Ybarra, and S. Benferhat, "From interval-valued probabilities to interval-valued possibilities: case studies of interval computation under constraints", *Proceedings of the 6th International Workshop on Reliable Engineering Computing REC'2014*, Chicago, Illinois, May 25–28, 2014, pp. 77–95.

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

[12] J. M. Mendel, *Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions*, Prentice-Hall, Upper Saddle River, 2001.

[13] J. M. Mendel and D. Wu, *Perceptual Computing: Aiding People in Making Subjective Judgments*, IEEE Press and Wiley, New York, 2010.

[14] H. T. Nguyen, V. Kreinovich, and Q. Zuo, "Interval-valued degrees of belief: applications of interval computations to expert systems and intelligent control", *International Journal of Uncertainty, Fuzziness, and Knowledge-Based Systems (IJUFKS)*, 1997, Vol. 5, No. 3, pp. 317–358.

[15] H. T. Nguyen and E. A. Walker, *A First Course in Fuzzy Logic*, Chapman and Hall/CRC, Boca Raton, Florida, 2006.

[16] N. R. Pal and J. C. Bezdek, "Measuring fuzzy uncertainty", *IEEE Transactions on Fuzzy Systems*, 1994, Vol. 2, No. 2, pp. 107–118.

[17] C. H. Papadimitriou, *Computational Complexity*, Addison Wesley, San Diego, 1994.

[18] G. Xiang, M. Ceberio, and V. Kreinovich, "Computing population variance and entropy under interval uncertainty: linear-time algorithms", *Reliable Computing*, 2007, Vol. 13, No. 6, pp. 467–488.

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