Relation Between Polling and Likert-Scale Approaches to Eliciting Membership Degrees Clarified by Quantum Computing

Renata Hax Sander Reiser,
Adriano Maron,
Lidiana Visintin
Universidade Federal de Pelotas – PPGC
96010-610, Pelotas, RS, Brazil
{reiser, akmaron, lvisintin}@inf.ufpel.edu.br

Ana Maria Abeijon
Universidade Católica de Pelotas
UCPel
96010-000, Pelotas, RS, Brazil
anabeijon@terra.com.br

Vladik Kreinovich
Department of Computer Science
University of Texas at El Paso
El Paso, TX 79968, USA
vladik@utep.edu

Abstract—**In fuzzy logic, there are two main approaches to eliciting membership degrees: an approach based on polling experts, and a Likert-scale approach, in which we ask experts to indicate their degree of certainty on a scale – e.g., on a scale form 0 to 10. Both approaches are reasonable, but they often lead to different membership degrees. In this paper, we analyze the relation between these two approaches, and we show that this relation can be made much clearer if we use models from quantum computing.**

I. Introduction

Need for fuzzy logic. A large part of our knowledge about the world is described in precise terms: we have equations (like Newton's equation) that describe the dynamics of systems, we have exact values characterizing the results of measuring physical quantities, we know the probabilities of different outcomes in a physical experiment, etc. However, a significant part of our knowledge comes from experts, be it medical doctors of skilled pilots. Experts can often only describe their knowledge by using imprecise ("fuzzy") words from natural language, such as "small", "young", etc. In contrast to well-defined terms, these words are not precise. While a medical doctor can be sure that a skin blemish of size 1 mm is small and a 5 cm size blemish is not small, this doctor may not be 100% certain whether intermediate values are small or not.

How can we describe this uncertainty? When we are absolutely confident in a statement, we declare this statement to be true; in a computer, "true" is usually represented as 1. When we are absolutely confident that a given statement is false, we declare this statement to be false; in a computer, "false" is usually represented as 0. To describe intermediate degrees of certainty, L. A. Zadeh proposed to use numbers between 0 and 1; see, e.g., [3], [5], [9]. These numbers are known as *membership degrees* because, e.g., the expert's degree of certainty that a 1 cm blemish is small can be viewed as a degree to which 1 cm belongs to the fuzzy set of all small values.

How can we elicit the corresponding degrees? There are many ways to elicit membership degrees from the experts.

One of these methods is *polling:* we ask several experts whether, e.g., a 1 cm blemish is small or not. If 7 out of 10 experts claim that it is small, we assign a degree $7/10 = 0.7$ to the statement that a 1 cm blemish is small. In general, if m out of n experts agree with the statement, we assign it a degree of certainty m/n.

When we only have one expert, we cannot use polling. In this case, we can ask the expert to mark his or her degree of certainty in this statement on a scale, e.g., from 0 to 10; such scales are known as *Likert scales*. If the expert selects 7 on a scale from 0 to 10, we assign, to this statement, a degree $7/10 = 0.7$. In general, if the expert describes his or her confidence in a statement by marking m on a scale from 0 to n, we assign, to this statement, a degree of certainty m/n.

Problem. Both above elicitation methods are reasonable, both lead to reasonable useful results. However, usually, these two methods led to different membership degrees. It is therefore reasonable to find out how these different degrees are connected.

Comment. Of course, degrees are subjective, so in general, different experts assign different Likert-scale degrees of certainty to the same statement. Because of this subjectivity, we cannot expect an exact one-to-one correspondence between the polling and Likert-scale degrees. What we want to discover is an *approximate* relation between the corresponding scales.

What we do in this paper. In this paper, we analyze the relation between the polling and Likert-scale elicitation techniques, and we show that this relation becomes clearer if we use formulas from quantum computing.

II. Relation Between Probabilistic (Polling) and Likert-Scale Elicitations: Analysis of the Problem

Probabilistic description of polling uncertainty. Formally, the formula m/n for the polling uncertainty is the same as the formula for a frequency (probability) of an event. This formal analogy makes sense. Indeed, our main objective in describing the expert's knowledge is to use it. For example, we want to know whether a 1 cm blemish is small or not because a medical expert describes her recommendations in terms of "small": one cure is proposed for a small blemish another for a large one. So, a possible way to find out whether a 1 cm blemish is small or not is to observe cases when treatment which works for small blemishes was actually used for a 1 cm blemish.

An expert who had such a patient and successfully used a cure intended for small blemishes (or, vice versa, unsuccessfully tried to use a cure intended for big blemishes) will vote that a 1 cm blemish is small. On the other hand, a doctor who, for a 1 cm blemish, unsuccessfully tried a cure intended for small blemishes (or, vice versa, successfully used a cure intended for big blemishes) will vote that a 1 cm blemish is not small.

In such an interpretation, the polling ratio m/n is equal to the frequency with which a randomly selected 1 cm blemish can be cured by a small-blemish cure. In general, in this interpretation, the polling ratio $\mu_P(x)$ describing to what extent a given value x satisfies the given property P is (approximately) equal to the frequency with which

- methods intended for objects satisfying the property P
- work for objects with the value x of the corresponding property.

From frequencies to a Likert scale: main idea. To determine the corresponding frequencies, we analyze several cases when a P-designed method was applied to x-valued objects. If out of N observations, the method worked M times, we take $\mu_P(x) = M/N$.

The resulting frequency values depend on how many situations we observed – and which exactly situations. For example, if on average, P-method works on a half on x-objects, it does not mean that we always get $\mu_P(x) = 1/2$: sometimes we may get $\mu_P(x) < 1/2$, sometimes $\mu_P(x) > 1/2$. This is similar to a usual statistical situation: if we flip two fair coins 10 times each, the first coin may falls heads 4 times and the second 6 times.

So, if for two different values x and x', we get different frequencies $f < f'$, we should not automatically conclude that objects with value x' are more certain to have property P: the observed difference in frequencies may be simply caused by the finiteness of the sample. The only case when we can make such a conclusion with confidence is when we are confident that the underlying probabilities p and p' differ. Such a conclusion is possible when the difference between the frequencies f and f' is sufficiently large.

In principle, out of N observations, we may have 0, 1, 2, …, N cases when the P-method worked. Usually, frequencies $0/N$ and $1/N$ may come from the same probability $p = p'$; similarly, $0/N$ and $2/N$ may come from the same probability, etc. As we increase $m = 1, 2, \ldots$, we will finally reach a value m_1 for which observed frequencies $f_0 = 0$ and $f_1 = m_1/N$ indicate that (with a given level of confidence) the actual (unknown) ing probabilities differ.

Similarly, the values $f_1 = m_1/N$ and $(m_1 + 1)/N$ are usually *indistinguishable*, in the sense that these two frequencies may result from the same probability. As we increase $m = m_1+1, m_1+2, \ldots$, we will reach the level m_2 for which observed frequencies $f_1 = m_1/N$ and $f_2 = m_2/N$ indicate that (with a given level of confidence) the corresponding probabilities differ, etc.

By repeating this procedure, we get a sequence of distinguishable frequencies $f_0 < f_1 < f_2 < \ldots$ such that each observed frequency is indistinguishable from one of the frequencies from this sequence:

- frequencies between f_0 and f_1 are indistinguishable from f_0;
- frequencies between f_1 and f_2 are indistinguishable from f_1;
- frequencies between f_2 and f_3 are indistinguishable from f_2;
- etc.

In other words, we have a finite number of distinguishable outcomes f_0, f_1, …, and the only reliable conclusion that we can make based on the observed frequency $f = M/N$ is to mark one of these outcomes. This is exactly what a Likert scale is about:

- we have a finite number of possible estimates, and
- to each situation, we place into correspondence one of these estimates.

From probabilities to a Likert scale: details. Let us use the known formulas from probability and statistics to derive explicit formulas for the corresponding frequencies f_0, f_1, …

Let us start with the simple case when we have a single observation. In this case:

- the probability of observing the favorable event (when a P-method worked on an x-object) is equal to p, and
- the probability of observing the opposite event (when a P-method did not work on an x-object) is equal to $1-p$.

In other words, the number X_1 of observed favorable events during a single observation is equal:

- to $X_1 = 1$ with probability p and
- to $X_1 = 0$ with probability $1 - p$.

Thus, the expected number $E_1 = E[X_1]$ of observed favorable events during a single observation is equal to $E_1 = p \cdot 1 + (1-p) \cdot 0 = p$. Similarly, the variance $\sigma_1^2 \stackrel{\text{def}}{=} E[(X_1 - E_1)^2]$ is equal to

$$\sigma_1^2 = p(1-p)^2 + (1-p)(0-p)^2 = p(1-p)^2 + (1-p)p^2 =$$

$$p(1-p)[p + (1-p)] = p(1-p); \qquad (1)$$

(see, e.g., [7]).

Let us now consider a more general situation, when we have $N \geq 1$ observations. If out of N observations, the favorable event happened in M cases, we compute the frequency f as the ratio $f = \dfrac{M}{N}$. One can easily check that this ratio is equal to the sum of N variables $X_1, \ldots, X_N$ corresponding to N observations divided by N:

$$f = \frac{X_1 + \ldots + X_N}{N}. \quad (2)$$

It is known that the expected value of the sum is equal to the sum of expected values [7], so

$$E[X_1 + \ldots + X_N] = E[X_1] + \ldots + E[X_N] = Np. \quad (3)$$

When a random variable is divided by a constant, its expected value decreases by the same constant [7], so we get

$$E[f] = \frac{E[X_1 + \ldots + X_N]}{N} = \frac{Np}{N} = p. \quad (4)$$

For N independent events, the variance of the sum is equal to the sum of variances [7], so we get

$$\sigma^2[X_1 + \ldots + X_N] = \sigma^2[X_1] + \ldots + \sigma^2[X_N] = Np(1-p). \quad (5)$$

When a random variable is divided by a constant, its variance decreases by the square of this constant [7], so we get

$$\sigma^2[f] = \frac{\sigma^2[X_1 + \ldots + X_N]}{N^2} = \frac{Np(1-p)}{N^2} = \frac{p(1-p)}{N}. \quad (6)$$

Due to the Central Limit Theorem [7], for large N, the distribution of the sum $X_1 + \ldots + X_N$ is close to Gaussian; thus, the distribution of the frequency $f = \dfrac{X_1 + \ldots + X_N}{N}$ is also close to Gaussian, with mean p and standard deviation

$$\sigma = \sqrt{\frac{p(1-p)}{N}}. \quad (7)$$

Since for large N, we have $f \approx p$, we can similarly conclude that for a given frequency f, the (unknown) value of the probability p is normally distributed, with mean f and standard deviation

$$\sigma = \sqrt{\frac{f(1-f)}{N}}. \quad (7a)$$

Now, we are ready to derive the formula for f_k. This derivation is based on the following situation:

- we have two different observations, with frequencies f and f', and
- we would like to know whether the observed difference $f \neq f'$ enables us to conclude that the corresponding probabilities p and p' are also different.

The corresponding distributions for p and p' are independent, so the difference $p' - p$ between the corresponding probabilities is equal to the difference between two independent (almost) normal distributions. Thus, the difference $p' - p$ is also normally distributed. The mean value of this difference is equal to the difference between the means, i.e., to $f' - f$, and the variance is equal to the sum of the variances, i.e., to

$$\sigma^2 = \frac{f(1-f)}{N} + \frac{f'(1-f')}{N}. \quad (8)$$

When can we guarantee – with a given degree of confidence – that $p' \neq p$? In statistics, it is known that the probability of being $\geq 2\sigma$ away from the mean E is $\approx 10\%$, the probability to be at least 3σ away from the mean is $\approx 0.1\%$, etc. For each degree of confidence, there is a value k_0 such that values from the interval $(E - k_0\sigma, E + k_0\sigma)$ are consistent with the observations, while values outside this interval are too improbable.

Thus, we can distinguish between probabilities when all consistent values (i.e., all values from the above interval) are different from 0, i.e., when $|E| \geq k_0\sigma$.

For each value f_k, what is the next distinguishable value f_{k+1}? According to the above description, it is the smallest value $f_{k+1} = f_k$ for which

$$\Delta f \stackrel{\text{def}}{=} |f_{k+1} - f_k| \geq k_0 \sqrt{\frac{f_k(1-f_k)}{N} + \frac{f_{k+1}(1-f_{k+1})}{N}}. \quad (9)$$

For large N, we get $f_{k+1} \approx f_k$, so

$$\frac{f_k(1-f_k)}{N} + \frac{f_{k+1}(1-f_{k+1})}{N} \approx 2\frac{f_k(1-f_k)}{N}, \quad (10)$$

and thus, the above condition (9) takes the form

$$\Delta f \geq k_0\sqrt{2}\sqrt{\frac{f_k(1-f_k)}{N}}. \quad (11)$$

The smallest of such values is

$$\Delta f = k_0\sqrt{2}\sqrt{\frac{f_k(1-f_k)}{N}}. \quad (12)$$

In the Likert-scale approach, frequencies $f_0 < f_1 < \ldots < f_k < f_{k+1} < \ldots$ correspond to marks 0, 1, …, k, $k+1$, …on the corresponding scale from 0 to n. According to the Likert-scale definition of the membership function, we assign to each mark k a value $\mu = k/n$.

Let us denote, by $\mu(f)$, the Likert-scale membership degree corresponding to the frequency f. Then, we have $\mu(f_k) = \dfrac{k}{n}$, and the next value $f_{k+1} = f_k + \Delta f$ corresponds to the degree $\mu(f_k + \Delta k) = \dfrac{k+1}{n}$. In other words, for every frequency f, we have

$$\mu(f + \Delta f) = \mu(f) + \frac{1}{n}. \quad (13)$$

When we have a reasonably large number of observations N, then the value Δf is small. When Δf is small, we have

$$\frac{\mu(f + \Delta f) - \mu(f)}{\Delta f} \approx \lim_{\Delta p \to 0} \frac{\mu(f + \Delta f) - \mu(f)}{\Delta f} = \mu'(f),$$

where $\mu'(f)$ denotes the derivative. Multiplying both sides of the approximate equality

$$\frac{\mu(f + \Delta f) - \mu(f)}{\Delta f} \approx \mu'(f) \quad (14)$$

by Δf, we conclude that

$$\mu(f + \Delta f) - \mu(f) \approx \mu'(f)\Delta f. \quad (15)$$

In this approximation, the formula $\mu(f+\Delta f)=\mu(f)+\dfrac{1}{n}$ takes the form $\mu'(f)\Delta f=\dfrac{1}{n}$, or, equivalently,

$$\mu'(f)=\frac{1}{n\Delta f}. \quad (16)$$

Substituting the above expression (12) for Δf into this formula, we conclude that

$$\mu'(f)=\frac{1}{nk_0\sqrt{2}\sqrt{\dfrac{f(1-f)}{N}}}. \quad (17)$$

We can simplify this expression into

$$\mu'(f)=\frac{c}{\sqrt{f(1-f)}} \quad (18)$$

for some constant c.

For $f=0$, when no expert claims that the statement is true, we should have $\mu(0)=0$. Thus, the general expression for $\mu(f)$ can be obtained by integrating the above formula:

$$\mu(f)=c\int_0^f \frac{dq}{\sqrt{q(1-q)}}. \quad (19)$$

A textbook way to compute this integral is to use an auxiliary variable t for which $q=\sin^2(t)$. In this case,

$$\sqrt{q}=\sin(t), \quad (20)$$

$$\sqrt{1-q}=\sqrt{1-\sin^2(t)}=\sqrt{\cos^2(t)}=\cos(t), \quad (21)$$

$$dq=d(\sin^2(t))=2\sin(t)\cos(t)\,dt, \quad (22)$$

and thus, the integral takes the form

$$\int_0^p \frac{dq}{\sqrt{q(1-q)}}=\int_0^{t_0}\frac{2\sin(t)\cos(t)\,dt}{\sin(t)\cos(t)}=$$

$$2\int_0^{t_0} dt=2t_0, \quad (23)$$

where t_0 is the value corresponding to f, i.e., the value for which $\sin^2(t_0)=f$. So, we conclude that $\mu(f)=2ct_0$, where $\sin^2(t_0)=f$. In other words, we have $t_0=C\mu$, where we denoted $C\stackrel{\text{def}}{=}\dfrac{1}{2c}$. Thus, the relation between μ and f takes the form

$$f=\sin^2(C\mu). \quad (24)$$

The value C can now be determined from the condition that the absolute confidence $\mu=1$ corresponds to $f=1$. For $\mu=1$, we get $f=1=\sin^2(C)$, so $C=\dfrac{\pi}{2}$. Finally, we get the relation between the polling membership value f and the Likert-scale membership value μ:

$$f\approx\sin^2\left(\frac{\pi}{2}\mu\right). \quad (25)$$

Discussion. At first glance, this relation looks very mathematical and non-intuitive. We will show, however, that it becomes much clearer if we use the techniques of quantum computing; see, e.g., [6].

III. Quantum Computing Clarifies the Relation between Polling and Likert-Type Degrees

Probabilities in quantum computing: reminder. The main idea of quantum computing is that by using quantum effects, we can often drastically speed up computations. For example, in classical physics, if we want to look for an element in an unsorted array of n elements, then we need at least n computational steps – because if we use fewer steps, we will not look into all n cells and thus, we may miss the desired element. Interestingly, in quantum case, we can perform the search in $\sqrt{n}$ steps (and $\sqrt{n}\ll n$); see, e.g., [1], [2], [6]. This possibility comes from the fact that in quantum physics, in addition to the usual classical states, we can also have *superpositions* of these states.

For a single bit, in addition to the usual states 0 and 1 – which are denoted as $\langle 0|$ and $\langle 1|$, we also have superpositions, i.e., states of the type $a_0\langle 0|+a_1\langle 1|$, where a_0 and a_1 are, in general, complex numbers (in quantum computing, only real values of a_0 and a_1 are used). Each such state can be described as a vector with coordinates (a_0,a_1) in a 2-D vector space. The corresponding quantum version of a bit is known as a *qubit*.

In a general qubit state $a_0\langle 0|+a_1\langle 1|$,

- the probability of observing 1 is equal to a_1^2, and
- the probability of observing 0 is a_0^2.

Since we always observe either 0 or 1, these probabilities must add up to 1, i.e., we must always have $a_0^2+a_1^2=1$. In geometric terms, this means that the vector (a_0,a_1) must be on the unit circle with a center at 0. Each such vector can be uniquely described by its angle φ with the axis corresponding to $\langle 0|$: in terms of this angle, we have $a_1=\sin(\varphi)$ and $a_0=\cos(\varphi)$.

Because of this relation, a state of a qubit is (almost) uniquely determined by the probability $p=a_1^2$ of observing 1. Once we know this probability, we can determine a_1 as $\pm\sqrt{p}$ (this $\pm$ is what we meant by almost uniquely), and we can determine a_0 as $\pm\sqrt{a_0^2}=\pm\sqrt{1-p}$.

Resulting relation between polling and Likert-scale degrees. For each probability p, we can form a qubit state

$$\sqrt{p}\,\langle 1|+\sqrt{1-p}\,\langle 0| \quad (26)$$

corresponding to this probability; see, e.g., see [4], [8]. For this state:

- on the one hand, due to the geometry of quantum states, we have $p=a_1^2=\sin^2(\varphi)$;
- on the other hand, due to the above relation (25) between frequencies and Likert-scale values, we have

$$p\approx f\approx\sin^2\left(\frac{\pi}{2}\mu\right). \quad (27)$$

Thus, we have $\sin^2(\varphi)\approx\sin^2\left(\dfrac{\pi}{2}\mu\right)$. In general, we would thus be able to conclude that $\sin(\varphi)\approx\pm\sin\left(\dfrac{\pi}{2}\mu\right)$. In our case, $\mu\in[0,1]$ and thus, $\dfrac{\pi}{2}\mu$ is between 0 and $\pi/2$, so

$\sin\left(\frac{\pi}{2}\mu\right) \geq 0$. The angle between the two straight lines is also usually defined as ranging from 0 to $\pi/2$. In particular, the angle φ between the straight line corresponding to the state $a_0\langle 0| + a_1\langle 1|$ and the axis corresponding to the "false" state $\langle 0|$ is also between 0 and $\pi/2$. Thus, $\sin(\varphi) \geq 0$ and therefore, $\sin(\varphi) \approx \sin\left(\frac{\pi}{2}\mu\right)$. Since both angles are from interval $[0, \pi/2]$, from the fact that they sines are approximately equal, we conclude that the angles are approximately equal as well, i.e., that the angle φ between the vector corresponding to this state and the vector corresponding to the "false" state $\langle 0|$ is equal to

$$\varphi \approx \frac{\pi}{2}\mu. \tag{28}$$

So, *the Likert-scale degree μ can be geometrically interpreted as (proportional to) the angle between the two states:*

$$\mu \approx \frac{2}{\pi}\varphi. \tag{29}$$

Fuzzy interpretation of a superposition between the two states (see [8]) **.** Superposition is a basic operation in quantum physics. In addition to superposition between the basic states $\langle 0|$ and $\langle 1|$, we can also consider a superposition of states

$$\sqrt{p}\,\langle 1| + \sqrt{1-p}\,\langle 0| \tag{30}$$

and

$$\sqrt{p'}\,\langle 1| + \sqrt{1-p'}\,\langle 0| \tag{31}$$

corresponding to uncertainty. To describe a superposition, we can simply add the corresponding vectors $(\sqrt{p}, \sqrt{1-p})$ and $(\sqrt{p'}, \sqrt{1-p'})$, and then "normalize" the resulting sum

$$(\sqrt{p} + \sqrt{p'}, \sqrt{1-p} + \sqrt{1-p'}), \tag{32}$$

i.e., divide it by the length

$$\sqrt{(\sqrt{p} + \sqrt{p'})^2 + (\sqrt{1-p} + \sqrt{1-p'})^2} \tag{33}$$

of this vector sum, to make sure that the resulting vector belongs to the unit circle (and is, thus, a legitimate quantum state). In terms of the probabilities p and p', the resulting vector has the form

$$(a''_1, a''_0) = \left(\frac{\sqrt{p} + \sqrt{p'}}{\sqrt{(\sqrt{p} + \sqrt{p'})^2 + (\sqrt{1-p} + \sqrt{1-p'})^2}}, \frac{\sqrt{1-p} + \sqrt{1-p'}}{\sqrt{(\sqrt{p} + \sqrt{p'})^2 + (\sqrt{1-p} + \sqrt{1-p'})^2}}\right) \tag{34}$$

with

$$a''_1 = \frac{\sqrt{p} + \sqrt{p'}}{\sqrt{(\sqrt{p} + \sqrt{p'})^2 + (\sqrt{1-p} + \sqrt{1-p'})^2}}. \tag{35}$$

Thus, the probability p'' of observing 1 in this state is equal to

$$p'' = (a''_1)^2 = \frac{(\sqrt{p} + \sqrt{p'})^2}{(\sqrt{p} + \sqrt{p'})^2 + (\sqrt{1-p} + \sqrt{1-p'})^2}. \tag{36}$$

In terms of probabilities, this looks like a very complex expression. However, in terms of angles, it becomes much simpler. Indeed, if we take a sum of two unit vectors at angles φ and φ' from the $\langle 0|$ axis, we get a bisecting vector at an angle

$$\varphi'' = \frac{\varphi + \varphi'}{2}. \tag{37}$$

Since the Likert-scale degree is simply proportional to the angle, we conclude that

$$\mu'' = \frac{\mu + \mu'}{2}. \tag{38}$$

So, *superposition corresponds to simple averaging of Likert-scale degrees.*

IV. Conclusion

In applications of fuzzy techniques, two main techniques are used for eliciting a membership degree μ of a given statement S:

- polling, when we ask several (n) experts and if m of them claim this statement to be true, we take $\mu = m/n$; and
- a Likert-scale approach, in which we ask an expert to mark his/her degree of certainty in the statement S on a scale from 0 to n; if the expert marks m, we take $\mu = m/n$.

Usually, these methods lead to different membership degrees. It is therefore reasonable to find out what is the relation between these two scales.

To uncover such a relation, we analyze the meaning of both scales. In both cases, we need to estimate the degree $\mu_P(x)$ to which the value x satisfies the given fuzzy property P: e.g., the degree to which a 1 cm skin blemish is small. A consequence of classifying the blemish to be small or not is whether we should apply, to this blemish, techniques designed for small blemishes. In practice, when we are not certain whether x satisfies the property P, this usually means that the corresponding techniques do not always work for objects of value x. By observing many such objects, we can find the probability p (to be more precise, frequency f) with which P-methods (i.e., methods intended for objects that satisfy the property P) work for x-objects (i.e., objects which have the value x of the corresponding quantity).

- An expert who observed that a P-method worked on an x-object will vote that x satisfies the property P.
- An expert who observed that a P-method did not work on an x-object will vote that x does not satisfy the property P.

Since the P-property works, on average, on the fraction p of x-objects, the proportion of experts who vote that x satisfies the property P is equal to the corresponding frequency f (and is, thus, approximately equal to the corresponding probability p). Thus, the polling membership degree is equal to the frequency f.

The same probabilities p can also explain different values on a Likert scale, if we take into account that the observed frequency f is, in general, slightly different from the actual probability p. For a sample of limited size N, nearby frequencies $f \approx f'$ can come from the same probability; only if the difference $f' - f$ is large enough, we can be sure (with a given degree of confidence) that the corresponding probabilities p and p' are also different. Thus, while potentially, we can have $N+1$ different frequencies 0, $1/N$, $2/N$, $\ldots$, $(N-1)/N$, 1, we have much fewer *distinguishable* ones, i.e., frequencies f and f' for which we are confident that they correspond to different probabilities. It is therefore natural to associate these distinguishable probabilities with marks on a Likert scale. This identification leads to the following relation between the polling membership value f and the Likert-scale membership value μ: $f \approx \sin^2\left(\frac{\pi}{2}\mu\right)$.

This relation is somewhat too mathematical and not very intuitively clear. It turns out that this relation becomes much clearer if we use models from quantum computing. Specifically, in quantum computing, an event with probability p is associated with a qubit state $a_0\langle 0| + a_1\langle 1|$ in which the probability of observing 1 is equal to p. It turns out that the corresponding value μ can be then geometrically interpreted as $\mu \approx \frac{2}{\pi}\varphi$, where φ is an angle between a straight line corresponding to the state $a_0\langle 0| + a_1\langle 1|$ and the axis corresponding to the "false" state $\langle 0|$. Thus, the use of quantum computing models clarifies the relation between the polling and Likert-scale membership degrees:

- a polling membership degree corresponds to the *probability* of observing a given property in a given state, while
- a Likert-scale membership degree is proportional to the *angle* between the given state and the "false" state (a state in which the given property is always false).

Acknowledgment

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, by Grants 1 T36 GM078000-01 and 1R43TR000173-01 from the National Institutes of Health, and by a grant on F-transforms from the Office of Naval Research.

Our sincere thanks to Regivan Santiago, Fernando Gomide, and to all organizers of the Brazilian Congress on Fuzzy Systems CBSF'2012 (Natal, Brazil, November 6–9, 2012) for valuable discussions which motivated this research. The authors are greatly thankful to the anonymous referees for important suggestions.

References

[1] L. K. Grover, "A fast quantum mechanical algorithm for database search", *Proc. 28th ACM Symp. on Theory of Computing*, 1996, pp. 212–219.

[2] L. K. Grover, "Quantum mechanics helps in searching for a needle in a haystack", *Phys. Rev. Lett.*, 1997, Vol. 79, No. 2, pp. 325–328.

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

[4] A. Maron, L. Visintin, R. H. S. Reiser, and M. Abeijon, "Interpreting Fuzzy Connectives from Quantum Computing – Case Study in Reichenbach Implication Class", *Proceedings of the Brazilian Congress on Fuzzy Systems CBSF'2012,* Natal, Brazil, November 6–9, 2012.

[5] H. T. Nguyen and E. A. Walker, *First Course In Fuzzy Logic*, CRC Press, Boca Raton, Florida, 2006.

[6] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge, U.K., 2000.

[7] D. J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, Chapman & Hall/CRC, Boca Raton, Florida, 2007.

[8] L. Visintin, A. Maron, R. H. S. Reiser, M. Abeijon, and V. Kreinovich, "Aggregation Operations from Quantum Computing", *Proceedings of the 2013 IEEE International Conference on Fuzzy Systems FUZZ-IEEE'2013*, Hyderabad, India, July 7–10, 2013.

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