3つのあいまいさ[†]

馬野 元秀*

今回は、ファジィ集合と確率と可能性という3つのあいまいさについて考えましょう*1。これらはいずれも区間 $[0, 1]$ の数値で表されるために、よく混乱してしまいます。

まず、普通の集合では、考えている対象 x が集合 A に属するか、属さないかのいずれか一方のみが成立します。属する場合を1、属さない場合を0として、 x から $\{0, 1\}$ への関数を考えると、これは定義関数 (defining function) または特徴関数 (characteristic function) と呼ばれているものになります (特性関数とも呼ばれます)*2。これにより集合と論理が結びつけられ、 $\{0, 1\}$ は2値論理の真理値と考えることができます。一方、ファジィ集合は L.A.Zadeh が導入したもので、考えている対象 x がファジィ集合 F に属する度合いを考えます。この度合いを区間 $[0, 1]$ の値とし、 x から区間 $[0, 1]$ への関数を考えると、メンバシップ関数 (membership function) となります。この区間 $[0, 1]$ はある種の無限多値論理の真理値と考えることができます。

一方、確率 (probability) は標本空間 (sample space) における事象 (event) を考えます。サイコロを振る場合では、「1の目が出る」、「2の目が出る」、…、「6の目が出る」が標本点または素事象 ω_i と呼ばれるもので、標本点全体が標本空間 Ω となります。試行により素事象 (標本点) のいずれか1つが出現します。事象 ω は標本空間の部分集合として定義され、標本空間全体の事象が起こる確率 $Pr(\Omega)$ を1とし、空の事象が起こる確率 $Pr(\emptyset)$ を0としますので、事象 ω が起こる確率 $Pr(\omega)$ は区間 $[0, 1]$ の値となります*3。

以上のように考えますと、ファジィ集合のメンバーシップ値と確率の違いは明確なように思われます。サイコロの場合で考えますと、「大きい目」や「小さい目」という概念は確率とは関係ありません。例えば、「大きい目」は $\{0.2/4, 0.7/5, 1/6\}$ のように 1～6 の目がファジィ集合「大きい目」にどれくらい属するかによっ

て決めることができます。これに事象は関係ありません。一方、確率では、標本空間 $\Omega = \{1, 2, 3, 4, 5, 6\}$ において、例えば、素事象がすべて同じ確率 $Pr(k) = 1/6$ ($k=1, 2, \dots, 6$) で起こると仮定すると、2以下の目が出る確率 $1/3$ や偶数の目が出る確率 $1/2$ などを計算することができます。

いまは、ファジィ集合「大きい目」のメンバシップ値は筆者が決めました。これは筆者の主観であって、合理的でないとよく言われます。しかし、この事情は確率でも同じです。サイコロの目のすべての素事象が同じ確率で起こると仮定しましたが、これはファジィ集合のメンバシップ値を決めるのと同じです。さらに、いまから振るサイコロがそのようなサイコロであるという保証也没有*4。ファジィ集合論ではメンバシップ関数が、確率論では素事象の確率が与えられたものとして議論を進めています*5。

ファジィ集合「大きい目」のメンバシップ関数を決めるのに、1人で決めると主観的になるので、多くの人に聞くという方法が考えられます。例えば、サイコロの目で5が大きいかどうかを多くの人に聞いて、1と0で答えてもらい、1の割合をメンバシップ値とする方法です。これは1と答える確率を表しており、同じ質問をしたときに1と答える人の割合を表わしています。しかし、1と答えた人が、5がどれくらい大きいと考えたかは分かりません。1と答えた人の割合とメンバシップ値に関係があるかもしれませんが、本来は別のものであると考えられます。

† Three Kinds of Ambiguities
Motohide UMANO

* 大阪府立大学 大学院理学系研究科 情報数理科学専攻
Department of Mathematics and Information Sciences,
Graduate School of Science, Osaka Prefecture University

*1 これら以外にも、あいまいさや不確かさが存在すると思われますが、ここでは、Zadehがさまざまなところで言及しているこれら3つについて考えます。

*2 日本数学会編「岩波数学辞典第3版」(1985年発行)による。2007年発行の「第4版」では、characteristic functionの訳から「特徴関数」がなくなって「特性関数」のみになり、「定義関数」の英語が indicator function になっています。

*3 確率を考えるには、これら以外にも加法性などの条件を考える必要があります。

*4 コインを投げる場合でも、表と裏のデザインの違いで、厳密には、1/2ずつにならないようです。また、試行を行なうと状況が大きく変化してしまい、多くの試行ができない場合もあります。

*5 同じように考えると、集合論でも論理でも具体的な値をどのように決めるかの部分の議論はされていません。集合や論理値が定義された後の処理方法が定式化されているだけです。

さて、3つ目の可能性について考えましょう。これは Zadeh が提案したもので*6、自然言語の文*7が持っている情報(意味)について考えたものです。次の3つの自然言語の文を考えてみましょう。

p_1 = 太郎は 24 歳である

p_2 = 太郎は 20 代である

p_3 = 太郎は若い

文 p_1 からは、太郎さんの年齢が 24 歳であることが分かります。これにはあいまいさもなく、問題ありません。では、文 p_2 はどうでしょうか。太郎さんの年齢が 20 歳以上 30 歳未満であることが分かりますが、この文から得られる情報「20 歳以上 30 歳未満」とは何でしょうか。

集合と考えて、太郎さんの年齢は区間 $[20, 30)$ のすべての年齢の集まりでしょうか*8。今の場合、太郎さんの年齢は区間 $[20, 30)$ のいずれかであって、 $[20, 30)$ 以外ではないと考えられます。これを Zadeh は太郎の年齢は区間 $[20, 30)$ のどの値である可能性もあるが、それ以外の値の可能性はないと考えました。そして、「可能性がある」を 1、「可能性がない」を 0 で表わし、可能性の分布の状態を表わす関数を定義しました。これは、集合の特徴関数の値 1 を「可能性がある」、0 を「可能性がない」とすると、集合の特徴関数と同じ形になります。そして、これを可能性分布関数 (possibility distribution function) と呼びました。

すると、文 p_1 からは、太郎の年齢に関する可能性分布が得られることになり、これを

$$\Pi_{\text{Age(太郎)}} = [20, 30)$$

と書きます。 $\Pi_{\text{Age(太郎)}}$ は「太郎」の「年齢 Age」の値が可能性分布になることを表わし、右辺はその可能性分布が $[20, 30)$ であることを表わしています。この場合、

区間 $[20, 30)$ にはファジィ的なあいまいさはありませんが、本当の値が区間の中のどこにあるか分からないという意味でのあいまいさがあることになります*9。

以上の議論で、可能性分布を考えなくても、等しい値をもつ確率分布で表現できると考えるかもしれませんが、しかし、確率分布を決めるには、確率に関する情報が必要です。いまの場合、区間 $[20, 30)$ という表現には確率の情報は入っていません。可能性のない年齢は起こり得ないので、区間 $[20, 30)$ 以外の年齢の確率は 0 ですが、 p_2 以外の情報がない状況では確率分布は「分からない」と考えるのが妥当だと思われます。もし $[20, 30)$ の日本人の年齢の分布が分かっていたら、太郎さんの年齢の確率分布を決める参考になると思われます。

さて、同じように考えると、命題 p_3 は

$$\Pi_{\text{Age(太郎)}} = \text{若い}$$

と表現することができます。この場合、「明らかに可能性がある」の 1 と「まったく可能性がない」の 0 の間の可能性の値を考え*10、ファジィ集合「若い」と同じ形の可能性分布を考えます*11。これが自然な拡張であると考えられます。この式は、属性にその可能な値を割り付けているので、「可能性割り当て方程式」(possibility assignment equation)と呼ばれますし、その属性に対してファジィな制約を付けているとも考えられますので、「弾力性のある制約」(elastic constraint)とも呼ばれています。

あいまいさを考えていくときに、確率だけではなく、ファジィ集合だけでもなく、可能性も考慮にいれて、3つのあいまいさを同時に考えていく必要があると考えられます。いずれも必要なあいまいさであると認識し、これらの統合理論を考えていくことが必要であると考えています。

[問い合わせ先]

〒599-8531 堺市中区学園町 1-1

大阪府立大学 大学院理学系研究科 情報数理学専攻

馬野 元秀

TEL: 072-254-9675

FAX: 072-254-9163

E-mail: umano@mi.s.osakafu-u.ac.jp

*6 “Fuzzy Sets as a Basis for a Theory of Possibility” という表題の論文で、雑誌 Fuzzy Sets and Systems の創刊号 (Vol.1, No.1, 1978) の最初の論文 (pp.3-28) でした。これを自然言語の意味表現に用いようと試みたのが、“PRUF – A Meaning Representation Language for Natural Languages” で、雑誌 International Journal of Man-Machine Studies の Vol.10, No.4 (1978) に掲載されました。

*7 Zadeh は命題 proposition と呼びましたが、ここでは文と呼ぶことにします。

*8 区間の表現で、 $[a, b)$ は a 以上で b 未満の半開区間を表わします。

*9 その意味で、可能性分布は多義性のあいまいさの表現方法と考えることができます。

*10 可能性でも、メンバシップ値でも、0.1 の差が何を意味し、具体的にどのように影響するかを議論することは難しいように思います。確率では、事象の発生に関連しているの、その解釈が可能であると思われます。

*11 Zadeh はメンバシップ値と可能性値の読み換えが可能であることを「可能性の仮定」(possibility postulate) と呼びました。