

解説



ファジィ集合に基づく自然言語の意味表現†

馬野元秀††

1. はじめに

自然言語には、さまざまなあいまいさが存在している。文法の解釈の多義性という意味でのあいまいさは、自然言語処理の分野では、非常によく研究されている¹⁾。しかし、自然言語の文がもつ意味のあいまいさについては、ほとんど研究されていない。

ファジィ集合²⁾は自然言語の単語（特に、形容詞）の意味を表現するために考え出されたということもあり、ファジィ集合により自然言語の意味的なあいまいさを表現しようという試みが、主に、L. A. Zadeh により行われている。その一つは、very や more or less などの副詞や、and や or などの接続詞をファジィ集合に対する演算子と考える方法³⁾である。そして、文については、PRUF⁴⁾として始まった研究が、現在では、テスト・スコア意味論^{5)~7)}としてまとめられている。

本稿では、あいまいな意味をもつ形容詞の意味表現から始めて、副詞、文、接続詞の意味表現について考え、PRUF とテスト・スコア意味論について説明する。

2. 自然言語の意味のあいまいさ

自然言語には、さまざまな意味的なあいまいさが含まれている。たとえば、自然言語の文

If a car which is offered for sale is cheap and much more than ten years old, then it probably is not in good shape. (1)

には、次のような5つの意味的なあいまいさが含まれている⁶⁾。すなわち、

(i) 時間に関する述語 much more than ten years old のファジィ的なあいまいさ

(ii) 値段に関する述語 cheap のファジィ的なあいまいさ

(iii) 形に関する述語 not in good shape のファジィ的なあいまいさ

(iv) 事象 the car is not in good shape に関する確率的なあいまいさ

(v) 確率 probable のファジィ的なあいまいさである。

実際の自然言語の文には、このようにさまざまなあいまいさが同時に含まれている。本解説では、主にファジィ的なあいまいさについてのみ考える。

ファジィ的なあいまいさを、「形容詞」、「副詞」、「文」、「接続詞」の4つに分けて考えよう。

(1) 形容詞

young, large, beautiful, attractive や much younger than などのように、その単語の意味が表す範囲がきちんと決められないという意味でのあいまいさである。式(1)の文では、(i)の much more than ten years old, (ii)の cheap, (iii)の in good shape, (v)の probable がこれに当たる。(v)の probable は確率値の表す範囲がきちんと決められないという意味である。

(2) 副詞

あいまいな意味をもつ形容詞を修飾して、その形容詞の意味を一定の規則で変形するもので、very, more or less, slightly, sort of, rather, pretty, not などがこれに当たる。また、先ほどの(i)では、more than ten years はあいまいではないが、much が付いて much more than ten years になるとあいまいになる^{*}。

(3) 文

文にはいろいろなものがあるが、ここでは主に

† Meaning Representation of Natural Language Based on Fuzzy Set Theory by Motohide UMANO (Department of Precision Engineering, Faculty of Engineering, Osaka University).

†† 大阪大学工学部精密工学教室

* あいまいさのない単語との比較においては、たとえば、more than ten years である度合いが11年と20年で同じかどうかという問題もある。

事実を表す文を考える。これを命題 (proposition) と呼ぶことにする^{*}。たとえば、

This building is large.

Joan is young.

などは命題の例である。

また、決まった真理値が付けられない命題がある。たとえば、

Show is white.

Trees are green.

Small cars are unsafe.

などは、普通は真であるが、いつも真であるとは限らない。このような命題を Zadeh は傾向命題 (disposition) と呼び、われわれの常識はこのような傾向命題の集合であると考えようと提案している⁹⁾。

(4) 接 続 詞

and, or, but などを用いて、形容詞や文をいろいろな形で接続する (文をつなぐ場合には, if-then も使うこともできる) ときに、あいまいさをどのように処理するかということを考える必要がある。形容詞をつなぐ場合 (形容詞が副詞で修飾されている場合も含むとする) の例としては, young or old, young but not very young や young and attractive などがある。また、文をつなぐ例としては、

Taro is old but Joan is young.

Snow is white and trees are green.

などがある。ただし、文をつなぐ場合には、接続詞自身には、あいまいさがないことが多い。また、式(1)の例は if-then を用いたものである。

3. 自然言語の意味表現

前章で述べた自然言語の意味はどのように表現すればよいのだろうか。順に考えていこう。

3.1 形容詞の意味表現

形容詞の意味はファジィ集合で表現される。これについては、いろいろなところで議論されているので、ここでは詳しくは述べない。ただ、ファジィ集合の決め方に2通りの方法があることだけを指摘しておこう。一つは young や large などのように、数値の属性を明確に決めることができる場合である。たとえば、young の場合には、まず

数値の年齢のうえでファジィ集合を定義し、これを用いて特定の人の年齢の young さを決めることができる。

もう一つは beautiful や attractive などのように、数値的な属性を決めることができない場合である。たとえば、beautiful の場合には、数値の属性がないので、特定の人の beautiful さは直接決めるしかない。このようなことは、主観を表現する単語に特に多い。

また、much younger than のように、二つの属性間の関係を表す場合には、ファジィ関係²⁾により表現する。

3.2 副詞の意味表現

副詞については、Zadeh はそれが修飾する形容詞のファジィ集合に対する演算子として定義した³⁾。このような意味から Zadeh はこれをファジィ修飾語 (fuzzy modifier) または言語修飾語 (linguistic hedge) と呼んだ。そして、G. Lakoff はいくつかの副詞の効果を図-1 のように示した⁸⁾。

図-1 のような各副詞のイメージを定式化する必要がある。実際、not については、補集合の演算として早くから定義されていた²⁾ が、very, more or less, slightly, sort of, rather, pretty は文献 3) で定義された。その後、Lakoff によりいくつかの定義が追加された⁸⁾。Zadeh と Lakoff の定義をまとめると次のようになる。

$$\text{not } A = \sim A \quad (2)$$

$$\text{very } A = \text{con}(A) = A^2 \quad (3)$$

$$\text{m. or l. } A = \text{dil}(A) = A^{0.5} \quad (4)$$

$$\text{slightly } A = \text{norm}(A \cap \sim \text{con}(A)) \quad (5)$$

$$\text{or} = \text{cint}(\text{norm}(plus(A) \cap \sim \text{con}(A))) \quad (6)$$

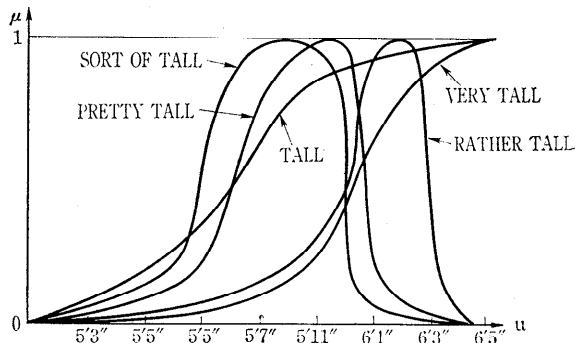


図-1 副詞の効果⁸⁾

^{*} 2値論理では、命題は真または偽であるかを判定できるものである。ファジィ論理では、適当な真理値を付けることができればよい。

$$\text{or} = \text{cin}(\text{norm}(\text{plus}(A) \cap \sim \text{plus}(\text{con}(A)))) \quad (7)$$

$$\text{sort of } A = \text{norm}(\sim \text{con}(\text{con}(A)) \cap \text{dil}(A)) \quad (8)$$

$$\text{rather } A = \text{cint}(\text{con}(A)) \quad (9)$$

$$\text{or} = \text{cint}(\text{con}(A)) \cap \sim \text{con}(A) \quad (10)$$

$$\text{pretty } A = \text{cint}(A) \cap \sim \text{cint}(\text{con}(A)) \quad (11)$$

ここで, m. or l. は more or less の略である. また, or は同じ副詞に対する別の定義を意味している(したがって, slightly には定義が三つあることになる). そして, ここで使われているファジィ集合の演算は, 次のように定義される.

$$\text{補集合 } C = \sim A : \mu_C(u) = 1 - \mu_A(u) \quad (12)$$

$$\text{べき乗 } C = A^r : \mu_C(u) = \mu_A(u)^r \quad (13)$$

$$\text{plus}(A) = A^{1.25} \quad (14)$$

$$\text{minus}(A) = A^{0.75} \quad (15)$$

$$\text{共通集合 } C = A \cap B : \mu_C(u) = \mu_A(u) \wedge \mu_B(u) \quad (16)$$

$$\text{対比強調 } C = \text{cint}(A) : \mu_C(u) = \begin{cases} 2 \cdot \mu_A(u)^2 & 0 \leq \mu_A(u) \leq 0.5 \\ 1 - 2(1 - \mu_A(u))^2 & 0.5 < \mu_A(u) \leq 1 \end{cases} \quad (17)$$

$$\text{正規化 } C = \text{norm}(A) : \mu_C(u) = \mu_A(u) \cdot \mu_A^* \quad (18)$$

(μ_A^* は最大値の逆数)

ここで, \wedge と \cdot は, それぞれ, min と掛け算を表している. なお, これらの副詞の定義は厳密なものではなくて, 第1次近似であるとしている. しかし, 上の定義は必ずしも図-1 のようにはならないことも報告されている⁹⁾.

3.3 文の意味表現

さて, ファジィ集合を使って文の意味をどのように表現するかについて説明しよう⁹⁾.

次の三つの自然言語の文(命題)の意味を考えてみよう.

$p_1 = \text{Jack is 20 years old}$

$p_2 = \text{Jack is a teen-ager}$

$p_3 = \text{Jack is young}$

命題 p_1 からは, Jack という人の年齢が20歳であることが分かる. これを

$\text{Age}(\text{Jack}) = 20$

と書くとする. すると, 命題 p_2 からは Jack の年齢が13歳以上20歳未満のいずれかの年齢であることが分かるので, これを

$\Pi_{\text{Age}(\text{Jack})} = [13, 20]$

と書くことにする. ここで, $\Pi_{\text{Age}(\text{Jack})}$ は Jack の年齢の可能な値を表している(Π は可能性 possibility の p に由来していると思われる). そして, $[13, 20]$ は13以上20未満の区間(すなわち, 集合)を表している. この場合, 区間 $[13, 20]$ は通常の集合であって, ファジィ集合ではないが, 本当の値がこの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.

同じように考えると, 命題 p_3 は

$\Pi_{\text{Age}(\text{Jack})} = \text{young}$

と表現することができる.

また, 副詞の付いた形容詞を含む命題

$p_4 = \text{Jack is not very young}$

は, 副詞の定義の式(2)と(3)を使って,

$\Pi_{\text{Age}(\text{Jack})} = \sim \text{con}(\text{young})$

と表現できる.

このような式は, 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) と呼ばれる. また, これは, その属性に対して制約を付けているとも考えられる. しかも, ファジィ集合による制約なので, 弾力性のある制約 (elastic constraint) とも呼ばれる.

次に, 傾向命題の意味表現について考えよう. 例として,

$dp = \text{Snow is white}$

を用いる. まず, この意味を, 命題

$p = \text{Most snow is white}$

であるとする*. この most をファジィ量限定子** (fuzzy quantifier) と呼んでいる. これは, 命題

$p' = 90\% \text{ of snow is white}$

のファジィ版と考えることができる. 命題 p' の

* 文献4)では, Usually snow is white であるとしているが, 考え方はほとんど同じである.

** most 以外の量限定子としては, たとえば, several, few, many, approximately n, close to n, a large number of などがある.

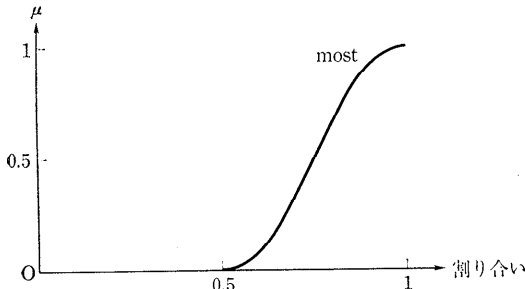


図-2 ファジィ集合 most

意味は

$$Prop(\text{white snow})=90\%$$

と考えることができる。ここで、 $Prop(\text{white snow})$ は雪のサンプル (あるいは、いままでに見た雪) のうちで、それが白色であった割合を表す属性である。すると、命題 p の意味は

$$\prod Prop(\text{white snow})=most$$

と考えることができる。ここで、 $most$ は割合のうえでのファジィ集合で、たとえば、図-2 のように定義できる。

白い雪の割合 $Prop(\text{white snow})$ は、すべての雪のサンプルとそのうちの白い雪の数の割合を計算すればよい。しかし、白い雪という概念はファジィ集合である。どのようにして、ファジィ集合の要素数を数えればよいのだろうか。これには、いくつかの方法があるが、シグマ・カウント (sigma-count) がよく使われる^{*}。これはメンバシップの度合いの算術和で定義される。いま、 A を $U = \{u_1, u_2, \dots, u_n\}$ におけるファジィ集合とすると、 A のシグマ・カウントは

$$\sum Count(A) = \sum_{i=1}^n \mu_A(u_i) \quad (19)$$

と定義される。通常の集合に対しては、この値は要素と数と等しくなる。これにより、白い雪の割合を計算することができる。

ところが、「雪である」という概念は必ずしも明確ではない (いわゆる「みぞれ」を考えれば、雪と雨との間に明確な境界はないと思われる)。また、

$$d = \text{Small cars are unsafe}$$

↓

$$p = \text{Most small cars are unsafe}$$

のような傾向命題や量限定子を含む命題を考える

^{*} これ以外の方法は、たとえば、文献 5) に述べられている。

と、主語の部分がファジィ集合の場合を考えておく必要がある。

そこで、一般の形として、

$$p = \text{Most } Ax \text{ are } B \quad (20)$$

を考えると、 A である x のうちの B である x の割合を計算する必要がある。これは

$$\frac{\sum Count(B/A)}{\sum Count(A)} = \frac{\sum Count(A \cap B)}{\sum Count(A)} \quad (21)$$

と定義することができる。ここで、 $A \cap B$ は A と B の共通集合で、式(16)で定義される。そして、この式(21)を相対シグマ・カウント (relative sigma-count) と呼ぶ。

これを使うと、式(20)の命題は、

$$p = Prop(Bx/Ax) \text{ is } most \quad (22)$$

となり、可能性割り当て方程式

$$\prod Prop(B(x)/A(x)) = most \quad (23)$$

として、表現できる。

3.4 接続詞の意味表現

and と or についてはファジィ集合が提案されたとき文献 2) に、共通集合と和集合の演算により定義されている。すなわち、

$$A \text{ and } B = A \cap B \quad (24)$$

$$A \text{ or } B = A \cup B \quad (25)$$

である。ここで、 $A \cap B$ は式(16)で定義され、 $C = A \cup B$ は

$$\mu_C(u) = \mu_A(u) \vee \mu_B(u) \quad (26)$$

と定義される。ここで、 \vee は max を表す。これを用いると、young or old や young but not very young などの同じ属性上の形容詞を接続したものはファジィ集合の演算に変換できる。ただし、最近では、and は T-ノルム (T-norm) と呼ばれる演算群に、or は T-コノルム (T-conorm) と呼ばれる演算群¹⁰⁾に対応しており、そのうちのどれを使うかはそのときの状況により決定するほうがよいと考えられている。また、平均演算子と呼ばれる and と or の中間の結果を与える演算子もある。

しかし、young and attractive のように異なる属性を接続している場合には、注意が必要である。この場合には、ファジィ関係が生成されると考えるべきであり、一般的に、接続詞 connect を使う場合には、 $C = A \text{ connect } B$ は

$$\mu_C(u, v) = \mu_A(u) * \mu_B(v) \quad (27)$$

となる。ここで、 $*$ は connect に対応した演算子で、たとえば、and に対しては min が、or 対

しては \max がよく使われる。

さらに、文に対しては、可能性割り当て方程式を考えるので、たとえば、

$p = \text{Joan is young and attractive}$

は、二つの属性を使って、

$$\begin{aligned} \Pi_{(Aoe(\text{Joan}), \text{Attractiveness}(\text{Joan}))} \\ = \text{young} \times \text{attractive} \end{aligned}$$

となる。ここで、 \times は直積 (Cartesian product) と呼ばれる演算で、

$$\mu_{A \times B}(u, v) = \mu_A(u) \wedge \mu_B(v) \quad (28)$$

と定義される。これは式(27)の演算 $*$ が \min の場合である。また、if x is A then y is B は、たとえば、

$$\Pi_{(x, y)} = \sim A \times V \oplus U \times B \quad (29)$$

と変換される。ただし、ファジィ制御ではこの式とは違った変換が使われることが多い。

さらに、傾向命題からどのように推論すればよいかという議論も行われている⁶⁾。

3.5 PRUF による文の意味表現

3.1 から 3.4 で述べてきた方法をまとめたものが PRUF である。PRUF は Possibilistic Relational Universal Fuzzy の頭文字をとったもので、それぞれ、「可能性的」、「関係的」、「万能的」、「ファジィ」という意味を表している。

PRUF は、自然言語の意味を表現するための言語で、自然言語の文の意味を可能性割り当て方程式で表現する。そして、自然言語の文を PRUF の表現に変換する規則を以下の4つの型に分けて示している(4つのうちの最後のものについては、前節までで説明していないが、ここでも詳細は省略する)。

(1) 修飾規則 (modification rule)

x を対象 (object) とし、 F をファジィ集合とすると、自然言語の文 x is F は、基本的には、 $\Pi_{A(x)} = F$ という PRUF の式に変換される。ここで、 $A(x)$ は x の属性を意味する。また、 x が n 変数のときは、 F はファジィ関係となる。そして、 F に not, very, more or less などの修飾語が付いた場合には、修飾語をファジィ集合に対する演算子と考えると、3.2 で述べたようにして計算する。

(2) 合成規則 (composition rule)

(1)の形の二つ以上の文を and, or, if-then を用いて結合したときの処理のしかたに関する規則で、要素の文に含まれているファジィ集合をもと

にし、3.3 で述べたようにしてファジィ関係をつくる。

(3) 量限定規則 (quantification rule)

この規則を適用する文の一般形は Qx are F である。 Q はファジィ量限定子と呼ばれるもので、most, many, few, some, almost などの単語に対応する。これは、ファジィ集合のグレードの和を求める関数やファジィ集合のうちである性質をもつものの比率を求める関数を使って、3.3 で述べたようにして変換する。

(4) 質限定規則 (qualification rule)

この規則を適用する文の一般形は x is F is τ である。そして、 τ は、true, not true, very true, false, more or less false などの真理値上のファジィ集合の場合(真理値限定と呼ばれる)と、probable, very probable, not probable, likely, very likely などの確率値上のファジィ集合の場合(確率限定と呼ばれる)と、possible, quite possible, almost impossible などの可能性の値のうえのファジィ集合の場合(可能性限定と呼ばれる)の3通りがある。これらの変換は、真理値限定の場合には F を τ で変形することにより、確率限定の場合にはファジィ事象の概念を利用することにより、可能性限定の場合には可能性測度を利用することより行う。これらの変換には、かなりの量の計算が必要な場合もある。

そして、文献4)では、(1)~(4)の例として次の英語の文の PRUF への変換の例が与えられている。

Ed is 30 years old.

Ed is young.

Ed is not very young.

Sally is very intelligent.

Edith is tall and blonde.

A man is tall.

All men are tall.

Most men are tall.

three tall men

several tall men

expensive red car with big trunk

John loves Pat.

John loves someone.

John loves everyone.

Someone loves someone.

Someone loves everyone.

Jill has many friends.

The man near the door is young.

Kent was walking slowly toward the door.

Herta is not very tall is very true.

Carole is very intelligent is very likely.

X is small is very true is likely.

men who are much taller than most men

Many men are much taller than most men.

Beth gave several big apples to each of her close friends.

4. テスト・スコア意味論による自然言語の意味表現

PRUF では、可能性を表すファジィ集合を適当な属性に割り当てることを、自然言語の意味と考えた。しかし、その後、Zadeh はこの考えをさらに進めて、テスト・スコア意味論 (test-score semantics) を提案した⁷⁾。

われわれが自然言語の文を受けとったとき、それが自分の知っている範囲の知識やデータとどれくらい合っているかを試そうとする。すなわち、自然言語の文に点数 (テスト・スコア) をつけようとするわけである。このためには、知識やデータのデータベースと点数を計算するための手続きが必要となる。もちろん、この手続きは受け取った自然言語の文に依存したものになる。そして、この手続き自身を受け取った自然言語の意味であると考えられるわけである。

例を与えよう。自然言語の文として、

$p = \text{Joan is young and attractive}$

を考える。そして、 p のテスト・スコアを計算するのに必要なデータベースを次のような関係とする。

$POPULATION [Name; Age; \mu_a]$

$+ YOUNG [Age; \mu]$

$POPULATION$ という名前の関係は三つ組の集合で、最初の要素は人の名前で、2番目の要素はその人の年齢で、3番目の要素はその人が魅力的である度合いである。また、関係 $YOUNG$ は順序対の集合で、最初の要素は属性 Age の値で、2番目の要素はその Age の値がファジィ述語 $young$ を満たす度合いである。この関係はファジィ述語 $young$ の意味を定義していると考えら

れる。

そして、このデータベースを用いてテスト・スコアを計算する手続きは、たとえば、次のようになる。

(1) $POPULATION$ において、属性 $Name$ の値が $Joan$ であるデータを検索して、 $Joan$ の年齢を求める。記号で書くと、

$$\begin{aligned} &Age(Joan) \\ &= \mu_{Age} POPULATION [Name = Joan] \end{aligned}$$

となる。この式で、 $rR[X=a]$ は、 R を $X=a$ に限定 (restrict) し、それを Y に射影 (project) することを表している。これは、 $X=a$ である組の Y の値を求めると考えればよい。

(2) $Joan$ が若い度合いを求める。これは

$$\tau_1 = \mu YOUNG [Age = Age(Joan)]$$

となる。これは $Joan$ の年齢の若さに関するテスト・スコアである。

(3) $Joan$ が魅力的である度合いを求める。これは

$$\tau_2 = \mu_a POPULATION [Name = Joan]$$

となる。これは $Joan$ の魅力に関するテスト・スコアである。

(4) 二つのテスト・スコア τ_1 と τ_2 をまとめて、全体的なテスト・スコアを計算する。このための集約演算子として、 \min 演算を使う。これは

$$\tau = \tau_1 \wedge \tau_2$$

となる。この全体的なテスト・スコアは、データベースのデータに対する $Joan$ is young and attractive の一致度を表している。

このとき、自然言語の文 p の意味は全体的なテスト・スコア τ によって表現されているのではなく、 τ を計算する手続きによって表現されていると考える。

もちろん、命題 p の意味表現、すなわち、手続きは、データベースを構成する関係に強く依存することになる。しかし、この関係は文を発する人も受け取る人も知っているとして仮定する (これがある言語を話す人間の間での共通の知識であると考えられる)。また、テスト・スコア意味論では、このデータベースは説明データベース (explanatory database) と呼ばれる。

次に、傾向命題の例として、

$d = \text{Snow is white}$

を考えよう。

意味表現の第1ステップは d にはない量限定子を復元することである。 d の意図した意味が、
命題

$p = \text{Most snow is white}$

であるとしよう。そして、データベースは、

$WHITE [Sample; \mu]$

$+ MOST [Proportion; \mu]$

であるとする。ここで、関係 $WHITE$ は雪の標本とそれが白色である度合い μ のリストである。一方、関係 $MOST$ は $Proportion$ の数値が $most$ の意味とどれくらい合っているかの度合いを定義している。

このとき、テスト手続きは、たとえば、次のようになる。

(1) 雪の標本 $S_i, i=1, 2, \dots, m$, の色が白である度合いを求める。これは、

$$\tau_i = \mu WHITE[Sample = S_i]$$

となる。

(2) 色が白の標本の割り合いを求める。これは、

$$\begin{aligned} \rho &= \frac{\sum Count(WHITE)}{m} \\ &= \frac{\tau_1 + \dots + \tau_m}{m} \end{aligned}$$

となる。

(3) ρ が $MOST$ による制約を満足する度合いを計算する。これは、

$$\tau = \mu MOST[Proportion = \rho]$$

となる。 τ は全体としてのテスト・スコアを表している。

このとき、 d の意味は τ を導くテストの手続きによって表されている。

以上をまとめると、命題 p の意味をテスト・スコア意味論により表現するには、次のようなステップで行えばよい。

(1) 説明データベースを構成するファジィ関係を選択する。

(2) 制約される属性 C_1, C_2, \dots, C_m を決定する。たとえば、命題 $Susan$ is young の場合には、制約される属性は「Susan の年齢」である。また、 $Most$ students are young の場合には、制約される属性は「学生のなかの若い学生の割り合い」である。

(3) 制約 $C_i, i=1, 2, \dots, m$, から部分的な

テスト・スコアを計算する。テストにより、各制約が満足される度合いを表すテスト・スコア τ_i を求める。普通、テスト・スコアは単位区間 $[0, 1]$ の一つの数値であるが、一般的には、単位区間上の確率分布または可能性分布となる。

(4) 部分的なテスト・スコア $\tau_1, \tau_2, \dots, \tau_m$ を少数のテスト・スコア $\bar{\tau}_1, \dots, \bar{\tau}_k$ に集約する。これは、

$$\tau = (\bar{\tau}_1, \dots, \bar{\tau}_k)$$

と表される。普通は、 $k=1$ で、スカラになる。これは、 p と説明データベースとの一致の度合いである。

上であげた例のほかに、文献 6), 7) では、次のような文に対するテスト手続きが書かれている。ただし、傾向命題については、右矢印の次にファジィ限定子を復元した命題を付記してある。

名詞句

many large balls

命題

Frenchmen are neither very tall nor very fat.

By far the richest man in France is bald.

傾向命題

Frenchmen are not very tall.

→Most Frenchmen are not very tall.

Young men like young women.

→Most young men like mostly young women.

Virginia smokes cigarettes.

→On the average Virginia smokes at least a few cigarettes a day.

Overeating causes obesity.

→Most of those who overeat are obese.

Heavy smoking causes lung cancer.

→The incidence of cases of lung cancer among heavy smokers is much higher than among those who are not heavy smokers.

Small families are friendly.

→In most small families almost all of the members are friendly with one another.

命令

Stay away from bald men.

ここでは、テスト・スコア意味論の基本的な考

え方を命題を用いて説明した。説明データベースやテスト手続きの作成は人間ならばそれほど難しいというわけではないが、計算機にやらせるのは非常に難しい。Zadeh も文献 7) で、「これは長期にわたる問題で、その完全な解決のためには、現時点のものよりも本質的に良い自然言語理解と知識表現法の開発を待たなければならない」と書いている。

5. おわりに

以上、Zadeh の研究を中心にして、あいまいな意味をもつ形容詞、副詞、文、接続詞の意味表現について考え、PRUF とテスト・スコア意味論について説明した。まだまだ不十分であるが、従来の自然言語の意味表現とは少し異なるアプローチであると思われる。また、現在のところ計算機上でのインプリメントは、残念ながらまったく行われていないが、インプリメントの経験を通しての改良が必要であると思われる。なお、ファジィ集合を用いて自然言語を扱うシステムのインプリメントに関する研究はほとんど行われていないが、筆者らの文献 11), 12) をあげておこう。

自然言語の研究は、非常に難しいが、人間の知能に深く関わっており、人間のもつあいまいさを避けて通ることができない分野でもある。本稿が、自然言語の研究者がファジィ理論について、ファジィ理論の研究者が自然言語について興味をもつきっかけになれば幸いである。

参考文献

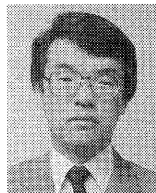
- 1) 長尾 確, 丸山 宏: 自然言語処理における曖昧さとその解消, 情報処理, Vol. 33, No. 7, pp. 746-756 (1992).
 - 2) Zadeh, L. A.: Fuzzy Sets, *Information and Control*, Vol. 8, pp. 338-353 (1965).
 - 3) Zadeh, L. A.: A Fuzzy-Set-Theoretic Interpretation of Linguistic Hedges, *Journal of Cybernetics*, Vol. 2, No. 2, pp. 4-34 (1972).
 - 4) Zadeh, L. A.: PRUF—A Meaning Representation Language for Natural Languages, *International Journal of Man-Machine Studies*, Vol. 10, pp. 395-460 (1978).
 - 5) Zadeh, L. A.: A Computational Approach to Fuzzy Quantifiers in Natural Languages, *Computers and Mathematics*, Vol. 9, pp. 149-184 (1983).
 - 6) Zadeh, L. A.: A Theory of Commonsense Knowledge, in Skala, H. J., Termini, S. and Trillas, E. (eds.): *Aspects of Vagueness*, Dordrecht: D. Reidel, pp. 257-296 (1984).
 - 7) Zadeh, L. A.: Test-Score Semantics as a Basis for a Computational Approach to the Representation of Meaning, *Literacy and Linguistic Computing*, Vol. 1, pp. 24-35 (1986).
 - 8) Lakoff, G.: Hedges: A Study in Meaning Criteria and the Logic of Fuzzy Concepts, *Journal of Philosophical Logic*, Vol. 2, pp. 458-508 (1973).
 - 9) 馬野元秀: ファジィ修飾語の効果の図的表示, 情報処理学会第 28 回 (昭和 59 年前期) 全国大会, pp. 1169-1170 (1984).
 - 10) 水本雅晴: 最近のファジィ理論, 情報処理, Vol. 29, No. 1, pp. 11-22 (1988).
 - 11) 馬野元秀: ファジィ集合の概念を用いた日本語質問応答システムの試作, 情報処理学会研究報告, Vol. 85, No. 36, AI 42-11 (知識工学と人工知能研究会), pp. 81-88 (1985).
 - 12) 馬野元秀: ファジィ集合の概念を用いた日本語質問応答システムの拡張—ファジィ推論の導入と一致度の処理法の追加—, 第 5 回ファジィシステムシンポジウム, pp. 433-438 (1980).
- なお, 最近, Yager, R. R., Ovchinnikov, S., Tong, R. M. and Nguyen, H. T. (eds.): *Fuzzy Sets and Applications—Selected Papers by L. A. Zadeh—*, John Wiley & Sons (1987).

の翻訳が

菅野道夫, 向殿政男 (監訳): ザデー・ファジィ理論, 日刊工業新聞社 (1992)

として, 出版された。これには, 文献 2), 3), 4), 5), 6), 7) の翻訳が掲載されている。

(平成 4 年 9 月 16 日受付)



馬野 元秀 (正会員)

1951 年生。1974 年, 大阪大学基礎工学部情報工学科卒業。1979 年, 同大学院後期課程修了。工学博士。

同年, 岡山理科大学理学部応用数学科講師。1985 年, 大阪大学大型計算機センター助手。

その後, 同センター講師, 助教授を経て, 1991 年より, 同大学工学部精密工学教室助教授。ファジィ集合論の応用, 特に, プログラミング言語, データベース, 知識情報処理への応用に関する研究に従事。電子情報通信学会, 日本ソフトウェア科学会, 人工知能学会, システム制御情報学会, 日本ファジィ学会, ACM, IFSA (International Fuzzy Systems Association) 各会員。