人　工　知　能 101－2
（1995．9．12）

論理構造上に定義された新しい確からしさの測度

藤本和則，　阿部明典，　松澤和光
NTT コミュニケーション科学研究所
〒238-03 横須賀市武 1-2356
tel: 0468-59-3117
email: {fujimoto, ave, matsu}@nttkb.ntt.jp

本稿では、問題の解候補集合が未知な系 (以下、不完全な系と呼ぶ) での推論の枠組みを提案する。まず、不完全な系での解候補の確からしさについて、我々の考えを述べ、その確からしさの測度 (下限確率と呼ぶ) を定義する。そして、下限確率の算出に必要な全ての知識を収集するのが困難であることを論じたあと、現実的に収集可能な知識のもとに下限確率を推定する計算法を提案する (この推定値を帰結度と呼ぶ)。そして最後に、帰結度に基づく推論例として、対話システムでの回答パターン決定問題を取り上げ、帰結度に基づいた決定法の有効性を示す。

A New Measure for Likelihood Defined with Logical Structures

Kazunori FUJIMOTO, Akinori ABE and Kazumitsu MATSUZAWA
NTT Communication Science Laboratories
1-2356 Take Yokosuka-shi Kanagawa 238-03 Japan
tel: +81-468-59-3117
email: {fujimoto, ave, matsu}@nttkb.ntt.jp

This paper proposes a new framework for reasoning with a measure for likelihood where the frame of descernment is unknown. First, our notion of likelihood for propositions in such incomplete conditions is described and a measure for the likelihood is defined (we call this measure Lower Probability). Difficulities in acquiring all knowledge for calculating Lower Probability are discussed and a new calculus for estimating this probability under acquireable knowledge is proposed. This paper also describes a methodology for applying this calculus to decision problems in dialogue systems.

1 はじめに

近年、人間と計算機の対話システムの研究においては、人間の発話から得られる情報の不確実性から、確率に基づいたシステム構築が注目されている [4]。しかしながら、確率は解候補が完全に網羅されていることを前提とするので、人間の自由な質問に対する解を見つけるといった“考えられる解候補が多い問題”への適用は難しい。

そこで、我々は、問題の解候補集合が未知な系 (以下、不完全な系と呼ぶ) での確からしさの測度「帰結度」に基づいて、確からしい解を選出する推論法 (アバウト論理 [1][2] と呼ぶ) の研究を進めている。こうした枠組みの実現は、対話システムに限らず不完全な知識を扱う AI において重要である。

不完全な系での確からしさの測度としては、空集合への確率割り当てを認める劣正規 DS 理論 [5][3] などが提案されているが、これらは、条件付確率や支持度などの数値データが個々の知識について獲得可能なことを前提にした理論である。対話システムのように膨大な量の知識を必要とする推論システム構築のためには、獲得が困難な数値データを必要としない確からしさ計算法の枠組みが望まれる。

そこで、我々は論理的知識を前提に、「結論を導くのに必要な事実のうちどの程度取得できたか」に着目して、各解候補の確からしさを表す帰結度を算出する計算法を提案した [1]。しかしながら、文献 [1] では帰結度の理論的意味付けが十分にされなかった。そこで、本稿では、不完全な系での確からしさについて我々の考えを述べ、なぜ帰結度がそうした確からしさを与えることができるかを中心に論じる。

本稿では、まず、不完全な系での「確率」の問題点を指摘し、次いで、そうした系での確からしさについて我々の捉え方を述べる。そして、不完全な系での確からしさと我々の提案する帰結度との関係を論じる。さらに、論理的知識のもとに帰結度の計算法を定式化する。そして最後に、帰結度に基づく推論例として、対話システムでの回答パターン決定問題を取り上げ、帰結度に基づいた決定法の有効性を示す。

2 不完全な系での確率値

確からしさの測度としては従来から確率測度がある。しかし、確率測度は、「全ての排反な解候補の確率値を足すと 1」という制約をもつので、用意すべき解候補に欠落があると、本来よりも大きな値を与えてしまう。

一般に排反で網羅的な n 個の解候補 $(C_1,\ldots,C_n)$ について、k 個の事実 $F=\{f^1,\ldots,f^k\}$ が得られたときの解候補 C_1 の事後確率値は、ベイズの定理より次のように書ける (ここに、$P(C_i)$ は解候補 C_i の事前確率、$P(f^m|C_i)$ は解候補 C_i に対する事実 f^m の条件付確率を表し、$C_*\in\{C_1,\ldots,C_n\}$ とする)。

$$P(C_*|F) = \frac{P(C_*)\prod_{m=1}^{k} P(f^m|C_*)}{\sum_{i=1}^{n}\{P(C_i)\prod_{m=1}^{k} P(f^m|C_i)\}} \qquad (1)$$

ここで、知識ベースに集める解候補に欠落があって、解候補 C_n に関する知識が得られなかったとする。このとき、式 (1) の分母は、i について 1 から $n-1$ までの和となるので、小さな値をとってしまう。その結果、$P(C_*|F)$ の値は、本来の値よりも大きな値として計算されることになる†。

このように確率値は、本来集めるべき解候補に欠落があれば確からしさを「過大評価」してしまい、逆に、集めるべきでない解候補を集めてしまえば「過小評価」してしまう。解候補に欠落のありえる知識ベースでは、確率値はその絶対的な値としての信頼性を失い、こうした信頼性の欠如は、確からしさに基づいた決定法での「判断の誤り」を引き起こす原因となってしまう。

3 不完全な系での確からしさ

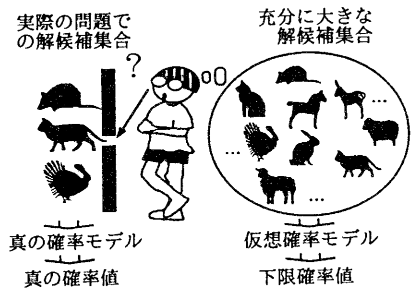

図 1: 不完全な系での確からしさ

ここでは、“不完全な系での確からしさ”について、我々の“確率論に基づいた捉え方”を述べる。

確率論によれば、実問題での解候補集合からなる確率モデルに基づいて算出された確率値が各解候補の確からしさを与える。一方、実問題での解候補集合を知ることができない問題では、確率論の計算原理を利用できないので、真の確からしさを知ることができない。

しかし人間は、実問題での解候補集合が分からなくても、直観的に各解候補の確からしさを得ることができる。つまり、一般に事実 F を観測したときの解候補 C の事後確率 $P(C|F)$ は、集めた解候補の数 n について単調減少するので、真の解候補集合を包含するほど充分に大きな N のもとでの $P(C|F)$ は、真の確率値の下限を保証することになる。我々は、人間はこうした“真の確率値の下限値”に着目して、不完全な系での確からしさを得ていると考えた (図 1)。

† “事前確率の和は 1” の制約より分母の $P(C_i)$ は相対的に大きくなるが、分子の $P(C_*)$ も同様に大きくなるのでこの効果は打ち消し合う

ここでは、こうした不完全な系での確からしさを下限確率値と呼び、次のように定義する。

定義 1 いま、事実 F が観測されたとする。このとき、ある解候補 C の“不完全な系での確からしさ”は、「充分に大きい N 個の解候補からなる確率モデルに基づいて算出された確率値 $P(C|F)$」によって与えられる。この $P(C|F)$ を C の下限確率値と呼ぶ。□

また、このような下限確率値を与える確率モデルを仮想確率モデルと呼ぶ。この下限確率値は、真の解候補集合での確率値の下限を保証するという意味で、不完全な系での確からしさを与えるとすることができる。

4 不完全な系での知識ベース構成法

推論システムにとって問題の解候補集合が未知であっても、以上のような“不完全な系での確からしさ”の考えに基づけば、各解候補の確からしさを下限確率値によって知ることができる。しかしながら、下限確率の算出に必要な仮想確率モデルは一般に膨大なので、それを全て収集するのは困難である。

仮想確率モデルの収集が困難である理由は、主に次の二つからなる。

- 仮想確率モデルでの解候補数 N は、その定義(実問題での解候補数より充分に大きい)から、一般に非常に大きな数となる。こうした解候補を全て集める必要があるので難しい。
- 確率モデルを記述するには、事前確率値と条件付確率値を $[0,1]$ の数値として用意する必要がある。こうした数値データは、論理的な二値データより集めにくいので難しい。

こうした理由から、知識ベースに仮想的確率モデルを用意しておくという設定は現実的とはいえない。

そこで、下限確率値を直接的に知るのをあきらめて、“仮想確率モデルから抽出可能な知識”のもとで、“下限確率値に近い値を推定する方法”を考える。すなわち、不完全な系での知識ベース構成の問題を「抽出コストの最小化」と「推定精度の最大化」のトレードオフに基づく問題(図 2)として捉える。

本稿では、こうしたトレードオフのもとでの「最適な抽出情報」について議論するには到っていないので、いささか直観的とはなるが、以下では、抽出コストが小さく、かつ、ある程度の推定精度が保証できるような知識ベースの構成法と下限確率値の推定法の一つを提案する。

具体的には、まず、仮想確率モデルから抽出する解候補を現実的に取得可能な n 個のみとし、さらに、それらの下限確率推定のための知識として、数値データではなく論理的な二値データのみを抽出して知識ベースを構成する。そして、そうした知識ベースをもとに、

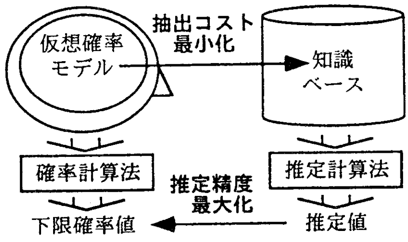

図 2: 不完全な系での知識ベース構成指針

(抽出した解候補については)下限確率値に近い値を推定できるようにしようという方法論の提案である。

4.1 下限確率値の推定原理

仮想確率モデルでは、各解候補の下限確率値 $P(C_*|F)$ は式(1)で与えられる(ただし式(1)での n を N に改める)。式(1)は、分子の項で分母分子を除すと次のように変形できる。

$$P(C_*|F) = \frac{1}{\sum_{i=1}^{N}\left\{\frac{P(C_i)}{P(C_*)}\prod_{m=1}^{k}\frac{P(f^m|C_i)}{P(f^m|C_*)}\right\}} \qquad (2)$$

式(2)から分かるように、全ての解候補 $C_i, i=1,\ldots,N$ にわたって、$\frac{P(f^m|C_i)}{P(f^m|C_*)}$ が小さいような事実 f^m が多く観測されればされるほど、下限確率値 $P(C_*|F)$ は大きくなる。この $\frac{P(f^m|C_i)}{P(f^m|C_*)}$ が全ての解候補にわたって小さいというのは、すなわち、事実 f^m が C_* にとって固有であるということを意味する。したがって、我々は、“各解候補について固有な事実”に着目して、「ある解候補に固有な事実が多く観測されれば、その下限確率値は大きな値をとる」とできるのではないかと考えた。

4.2 仮想確率モデルからの抽出情報

以下では、仮想確率モデルから抽出する解候補は、現実的に取得可能な解候補 $C_i, i=1,\ldots,n(<N)$ のみであることを前提とする。

我々は、4.1 節での推定原理に基づいて、各解候補 C_i の下限確率値を推定するために必要な情報を次のように分類した(抽出が容易なものから順に並べた)。

1. 各 C_i にとって、事実 f^m が「固有か否か」の二値情報
2. 各 C_i について、下限確率値を 1 に近い値にする固有な事実 f^m の集合情報
3. 各 C_i にとって、事実 f^m が「固有な程度」の数値情報
4. 2. と 3. を合わせた情報

1. では、推定値の精度を確保することが困難であることが予想され、3.,4. では、各固有の程度を定量的に取得する必要が生じるため、抽出のコストが飛躍的に大きくなる。したがって、ここでは、2. の情報を抽出した知識ベースを考え、その情報のもとで下限確率値を推定する計算法を構築する。

2. の情報とは、すなわち、“各解候補 C_i を「真に限りなく近い」とするのに最小限必要な事実の集合”という単純な論理的構造である。したがって、仮想確率モデルから 2. のみの情報を抽出するということは、各解候補 C_i について、論理的な帰結情報のみを抽出し、それ以外の数値データ等は一切抽出しないことを意味する (図 3)。

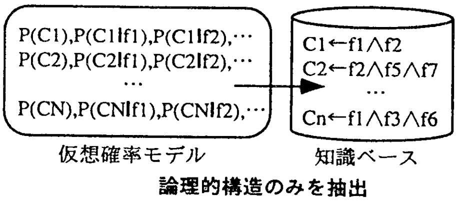

図 3: 仮想確率モデルからの抽出情報

ここでは、$C_i,\ldots,C_n$ について、こうした比較的単純な知識しか得られていなくても、次のような三つの性質に基づいて、下限確率値に近い値を推定できると考える。

- 固有な事実が全く観測されなければ下限確率値は 0 に近くなる
- 固有な事実が過不足なく全て観測されれば下限確率値は 1 に近くなる
- 固有な事実の観測数が増えれば下限確率値は増加する

以下では、各 C_i についての“下限確率値を 1 に近い値にする固有な事実 f_i^m の集合”を確率的極小支持集合 (以下、PMS) と呼び、次のように定義する。

定義 2 仮想確率モデルにおいて、$P(C_i|H)\geq 1-\varepsilon$ (ε は 0 に近い微小な正数) を満たす極小の事実集合 $H=\{f_i^1,\ldots,f_i^r\}$ を C_i の確率的極小支持集合 *(Probabilistic Minimal Supports)* と呼び、$C_i \leftarrow f_i^1\wedge\cdots\wedge f_i^r$ と書く。□

すなわち、PMS に含まれる事実が過不足なく観測された解候補は、下限確率値が $1-\varepsilon$ 以上であることが保証される。

以下では、下限確率値の推定値を帰結度と呼び、解候補 $C_1,\ldots,C_n$ についての PMS 式集合からなる知識ベースにおいて、各解候補の帰結度を与える関数の定式化を行う。

5 帰結度の計算法

ここでは、事実集合 $F=\{f_*^1,\ldots,f_*^k\}$ が観測されたときの解候補 C_* の帰結度を与える関数 $E:(C,F)\rightarrow[0,1]$ の定式化を行う (今回は、ある解候補 C_i の PMS が事実集合 F を包含している場合のみを対象とする)。

5.1 帰結度の満たすべき条件

いま、PMS が H の解候補 C について、事実集合 F が観測されたとする。ここでは、このときの帰結度関数 E の満たすべき条件として、次の三つを抽出した。

最小条件 H 中の事実が一つも観測されなければ最小値 0 を与える ($H\cap F=\phi\Rightarrow E(C,F)=0$)。

最大条件 H 中の事実が全て観測されれば最大値 1 を与える ($H\subseteq F\Rightarrow E(C,F)=1$)。

増加性条件 H 中の事実が多く観測されればされるほど帰結度は増加する
($H\cap F_1\subseteq H\cap F_2\Rightarrow E(C,F_1)\leq E(C,F_2)$, 但し等号は $H\cap F_1=H\cap F_2$ のときに限る)。

5.2 帰結度計算の定式化

以上の三つの条件に基づいて、ここでは、帰結度を与える関数を次のように定義する。

定義 3 PMS が $\{f^1,\ldots,f^r\}$ の解候補 C について、事実集合 F が観測されたとき、次式で定められる関数 $E:(C,F)\rightarrow[0,1]$ によって与えられる推定値を帰結度と呼ぶ。

$$E(C,F)=\sum_{m=1}^{r}\alpha_m T(f^m) \tag{3}$$

ここに、T は $f^m\in F$ なら 1, $f^m\notin F$ なら 0 を与える関数、各事実に付けられた重み α_m は、$\sum_{m=1}^{r}\alpha_m=1$ なる正数である。□

式 (3) での重み α_m は、C にとって f^m が固有な程度を表す。これは、$P(f^m|C_i), i=1,\ldots,N$ から定められる値ではあるが、我々は、こうした数値データがなくても、「少数の PMS にしか含まれない事実は、その事実を含む解候補にとって固有である」という傾向に着目して、各事実の固有程度を知ることができるのではないかと考えている (但し、ある程度の規模の PMS 式集合が得られていることを前提とする)。この可能性に関しては別途報告することにする (知識ベースの論理的構造のみから α_m を算出する計算例については、文献 [1] を参照されたい)。

事実の集合が観測されたとき、式 (3) に基づいて、解候補の下限確率値の推定値を帰結度として算出することができる。

6 PMS 抽出と帰結度計算の例

ここでは、簡単な仮想確率モデルをもとに、仮想モデルからの PMS の抽出と、帰結度の計算の例を示す。

6.1 仮想確率モデル

仮想確率モデルとして、20 個の解候補 C_i と、100 個の事実 $F\{f_1^1,\ldots,f_{20}^5\}$ についての確率モデルを仮定する。簡単のため、事前確率値は各 C_i について均一とする。また、各 C_i に固有な事実がそれぞれ 5 個 $(f_i^1,\ldots,f_i^5)$ 存在するとし、条件付確率は、C_i について固有である事実 f_i^m では $P(f_i^m|C_i) = 0.8$、そうでない事実 f_j^m については $P(f_j^m|C_i) = 0.2$ とする (以上のような確率モデルは確率的に完全である)。

- $\forall i, P(C_i) = 1/N = 0.05$
- $\forall i, P(f_i^m|C_i) = 0.8, m = 1,\ldots,5$
- $\forall i \neq j, P(f_j^m|C_i) = 0.2, m = 1,\ldots,5$

6.2 PMS の抽出

各 C_i について PMS は、定義 2 に基づいて仮想確率モデルから抽出する (ここでは定義 2 での ε を 0.05 とする、また事実集合 F が観測されたときの C_i の下限確率値は、式 (1) に仮想確率モデルのデータを代入して計算できる)。

各 C_i にとって固有な事実が 4 個しか観測されなかったときは、下限確率値は 0.93 となり 1-0.05=0.95 を越えず、固有な事実が 5 個観測されたときには、下限確率値は 0.98 となり初めて 0.95 を越えることがわかる。よって、各 C_i に固有な 5 個の事実からなる集合がそれぞれの C_i にとって唯一の PMS として抽出される。

以上のような手続きによって、次のような $n(\leq N)$ 個の PMS 式を得ることができる。

$$C_i \leftarrow f_i^1 \wedge \cdots \wedge f_i^5, i = 1,\ldots,n \tag{4}$$

6.3 帰結度計算

ここでは、ある事実集合 F が観測されたとき、"仮想確率モデルから下限確率値"、"PMS 式集合から帰結度" をそれぞれ計算し、それらの関係を調べる (今回は、帰結度は式 (3) で各事実の重みが等しいとして計算する)。

観測された事実の個数が 0 個から 5 個のそれぞれの場合について、帰結度と下限確率値の計算結果を表 1 に示す。知識ベース内には、仮想確率モデル中の一部の解候補しか抽出せず、かつ、数値的な確率データが得られてないにもかかわらず、帰結度計算によって、(仮想確率モデル上で得られる) 下限確率値におおむね近い値を推定できることがわかる。

表 1: 帰結度と下限確率値

事実 F	帰結度 $E(C_*, F)$	下限確率値 $P(C_*\|F)$
なし	0.0	0.050
$\{f_*^1\}$	0.2	0.174
$\{f_*^1, f_*^2\}$	0.4	0.457
$\{f_*^1,\ldots,f_*^3\}$	0.6	0.771
$\{f_*^1,\ldots,f_*^4\}$	0.8	0.931
$\{f_*^1,\ldots,f_*^5\}$	1.0	0.982

7 回答パターン決定問題への適用

以上では、不完全な系での確からしさを推定する帰結度の計算法を定式化した。本章では、こうした帰結度に基づいて、対話システムでの "回答パターン" を決定する方法を提案する。扱う問題としては、人間が観測事実をシステムに投入し、システムがその観測事実をもとに対象の属するカテゴリを回答するという対話型分類問題を取り上げる (図 4)。

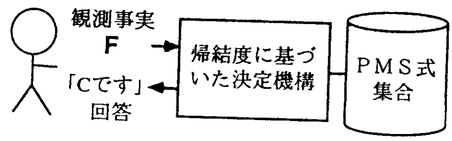

図 4: 対話型分類問題

以下では、"得られる利益の期待値 (期待利益と呼ぶ)" に着目して提案の手法と従来の確率値に基づく手法を比較し、提案の決定法の有効性を確認する。

7.1 帰結度に基づく決定法

知識ベースに集められた解候補に欠落があり得る場合には、人間の質問に対する解がシステム内に存在するとは限らない。したがって、「解がわからない」という回答を許す必要がある。ここでは、帰結度を用いた回答パターン決定法として、次のような決定法を提案する。

帰結度に基づく決定法 いま、事実 F が観測されたとする。このとき最大の帰結度 $E(C,F)$ が、あらかじめ定められた閾値 $D(0\leq D\leq 1)$ 以上であれば「解は C である」と決定し、未満であれば「解は分からない」と決定する。

一方、確率値に基づく決定法 (例えばベイズ決定 [6]) は、真の解候補が明らかでない問題においては、(確

率値が集めた個数に依存するので)上述のような絶対的な閾値に基づいた決定は意味をもたない。これに対して帰結度は、集めた解候補の個数に依存しないので、上述のような決定法が可能となるのである。

7.2 帰結度に基づく決定法の評価

ここでは、得られる期待利益に着目して、帰結度に基づく決定法の評価を行う。まず、6.2 節に示した PMS 式(4)を仮定する(6 章での帰結度計算と同様に各事実の重みは均等とする)。そして、決定の結果得られる利益を、回答した解が正解なら 1 点、回答しなければ 0 点、回答した解が不正解なら -1 点と定める。

いま、仮に閾値 D=0.5 とすると、提案の決定法に基づき、$E(C,F)>0.5$、すなわち、得られた事実が 3 個以上なら「解は C である」、$E(C,F)\leq 0.5$、すなわち、得られた事実が 3 個未満なら「解は分からない」と決定する。このとき、得られた事実の個数を k 個(k は $0\leq k\leq 5$ の整数)とし、k は $0,\ldots,5$ を等確率(1/6)に取りうるとすると、得られる利益の期待値は次式で与えられる。

$$EP=\frac{1}{6}\left[2\sum_{k=3}^{5}\frac{\prod_{m=1}^{k}P(f^m|C_*)}{\sum_{i=1}^{20}\{\prod_{m=1}^{k}P(f^m|C_i)\}}-1\right] \quad (5)$$

これを図 5 に示す。図 5 において、横軸は真の解候補の個数 n、縦軸は得られる利益の期待値 EP を表す。比較のため、確率値最大の結論を答える決定法の得点期待値を図 5 中に示す。このように、確率値最大の解を答える決定法では、期待利益が真の解候補の個数に大きく依存してしまうのに対して、提案の決定法では、確からしさが保証されたものを選出して答えるので、安定した期待利益を得ることができる。

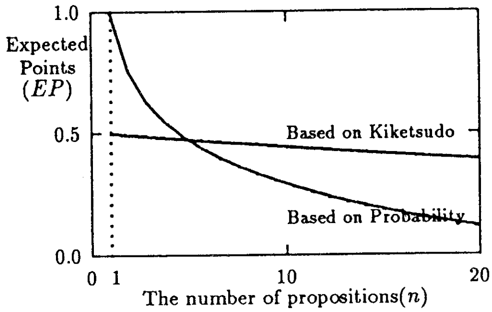

図 5: 得点の期待値

8 おわりに

本稿では、問題の解候補集合が未知な系での推論の枠組みを提案した。まず、不完全な系での解候補の確からしさは、充分に大きな解候補集合での確率値で与えられるとし、その確率値を下限確率値として定義した。そして、下限確率値の算出に必要な全ての知識の収集は、「解候補の数が膨大」、「数値データの抽出コストが大きい」という二つの観点から困難であることを論じた。そこで、各解候補について、“下限確率値を 1 に近くする事実の集合(確率的極小支持集合)”という論理的構造のみを知識として収集し、そうした知識のもとに下限確率値を推定する計算法を提案した(この推定値を帰結度と呼んだ)。

さらに、帰結度に基づく推論例として、対話システムでの回答パターン決定問題を取り上げ、帰結度を用いて適切な回答を決定する方法を提案した。そして、不完全な系において、真の解候補の個数への依存の少ない「安定した期待利益」を得ることが可能になることを示した。

今後の課題としては、“帰結度計算の定式化の精錬”と、“確率的極小支持集合からなる知識のもつ性質の解析”をあげることができる。後者に関しては、ホーン節集合との類似性を検討中である。これらについては、稿を改めて報告したい。

参考文献

[1] 藤本他, “ルールの論理的構造に着目した結論の“尤もらしさ”の定量化”, 情報処理学会論文誌, Vol.36, No.8,pp.2071-2074,1995.

[2] 松澤他, “アバウト推論方式の基本構想について”, 電子情報通信学会, 信学技報 AI93-77(1994-01),pp.41-48,1994.

[3] 菅野 道夫, 室伏 俊明, “ファジィ測度”, 日本ファジィ学会編, 日刊工業新聞社, 1994.

[4] 秋葉 友良, 田中 穂積, “ベイジアンネットワークを用いた対話システム: ユーザモデルの推定”, 人工知能学会, 人工知能研究会 SIG-SLUD-9303-2,pp.9-16,1994.

[5] Shafer,G., “A mathematical theory of evidence”, Princeton, 1976.

[6] Richard O. Duda, Peter E. Hart, “Pattern Classification and Scene Analysis”, John Wiley and Sons,1973.