

曖昧性解消のための優先度計算法：  
優先度 = 意味ファジィ  $\wedge$  意味確率  $\wedge$  構文確率

李 航

NEC C&C 情報研究所  
〒216 川崎市宮前区宮崎 4-1-1  
E-mail: lihang@ibl.cl.nec.co.jp

本稿では自然言語解析における曖昧性解消のための優先度の計算法を提案する。本研究では、優先度に影響を与える構文的、意味的な要因を総合的に考察し、確率的なアプローチで優先度を計算する方法を考案した。具体的には意味ファジィと意味確率と構文確率のベクトルを優先度とする。意味ファジィは解釈の格フレームの意味的な整合性を表し、意味確率は解釈の格フレームにおける単語の共起頻度を表し、構文確率は解釈の句構造の出現頻度を表す。この優先度の計算法は、信頼性が高く、実用的である。

**A Method of Computing Preference for Disambiguation:  
Preference = Semantic Fuzziness  $\wedge$  Semantic Probability  
 $\wedge$  Syntactic Probability**

Hang LI

C&C Information Technology Research Laboratories, NEC  
Miyazaki 4-1-1, Miyamae-ku, Kawasaki 216, Japan  
E-mail: lihang@ibl.cl.nec.co.jp

This paper proposes a method of computing preference for disambiguation in natural language analysis. Preference is modeled as a vector of semantic fuzziness, semantic probability and syntactic probability. Semantic fuzziness represents the degree of semantic consistency of a given interpretation. Semantic probability represents the co-occurrence frequency between all the words in a given interpretation. Syntactic probability represents the occurrence frequency of the syntactic structure of a given interpretation. Using this preference method, disambiguation can be achieved more satisfactorily.

## 1 はじめに

自然言語を解析することは自然言語を一意的に解釈することである。一般的に自然言語の解析は構文解析と意味解析と文脈処理の過程から構成される(図1)。



図1: 自然言語解析のモデル

構文解析では自然言語の解釈の文法的な構造を明らかにする。意味解析では構文的に妥当とされる解釈の中から意味的に異常な解釈を排除する<sup>1</sup>。自然言語を構文意味解析した結果、多くの場合複数の解釈が得られる。例えば、

I saw a girl with a telescope. (1)

を構文意味解析した結果、「私は望遠鏡で女の子を見た」と「私は望遠鏡を持っている女の子を見た」という二通りの解釈が得られる。文脈処理では文脈知識を利用して構文意味解析で得られる複数の解釈の一つの解釈に特定する。例えば、後に

She bought it to see the comet. (2)

という文が現れたら、前文の解釈が「私は望遠鏡を持っている女の子を見た」であると判断していいであろう。しかし、現在の技術レベルでは本格的な文脈処理が実現しにくい。文脈処理を含まない自然言語解析では、以上のような曖昧性が解消できず、自然言語解析のもっとも大きな問題になっている。

一方、人間は文脈知識がない場合でも自然言語の複数の解釈の存在を感じながら、もっともらしい解釈をほぼ一意的に与えてしまう。人間が「構文意味解析」を行った後に得られた複数の解釈の「優先度」を計算しているように思われる。人間と同じように構文意味解析の過程で優先度を計算する戦略が早くから提案された[辻井86]。具体的には、構文意味解析で得られた各解釈の優先度<sup>2</sup>を計算し、文脈処理があれば、文脈処理によって解釈を決定し、文脈処理がなければ、優先度のもっとも高い解釈を入力解釈とする。いうまでもなく、この戦略では、優先度の計算方法が重要な課題になる。

従来では、優先度はヒューリスティクス、ニューラルネットワーク、確率によって計算されている。ヒューリスティクスによる優先度は理論的な保証が欠けているため、信頼性が低い。ニューラルネットワークによる優先度は、規模の大きいシステムでは膨大な学習時間が必要なので、現時点では実用的とはいえない。確率による優先度は、数学的な基礎がしっかりしているし、システムの規模が大きくなっても統計がで

<sup>1</sup>従来の意味解析では意味的に異常な解釈を排除する。しかし、後に述べるように文脈によっては意味的に異常な解釈でも解釈可能であるので、意味解析では意味的に異常な解釈も残し、それぞれの解釈の解釈される可能性を計算すべきである。

<sup>2</sup>本研究でいっている優先度は文脈的要素を含まない。

きるのも、もっとも実用的といえる。しかし、従来の研究では、解釈の句構造の確率を優先度としたり、或いは解釈における単語の共起確率を優先度としたりしている。優先度に影響を与える構文的、意味的な要因を総合的に考察し、確率的なアプローチで優先度を計算する研究はまだない。

本研究では、自然言語の解釈の優先度を以下のように計算する。

$$\text{優先度} = [\text{意味ファジィ}, \text{意味確率}, \text{構文確率}] \quad (3)$$

つまり、意味ファジィと意味確率と構文確率のベクトルを優先度とする。意味ファジィは解釈の格フレームの意味的な整合性を表し、意味確率は解釈の格フレームにおける単語の共起頻度を表し、構文確率は解釈の句構造の出現頻度を表す。この優先度の計算法は、信頼性が高く、実用的である。

本稿の構成は以下のとおりである。まず、第2章で従来の研究とその問題点について述べる。第3章では、本研究提案する優先度の計算法について述べる。第4章でまとめる。また、付録で主観確率による意味的な優先度について考察する。

## 2 従来の研究と問題点

この章では、現在までの優先度に関する研究を簡単に紹介する。

### 2.1 言語学からみる優先度

FrazierとFodorは英語の優先度に関する原理 Minimal Attachment PrincipleとRight Association Principleを提唱した[Frazier79]。MAPによると、修飾語(Modifier)が項(Argument)になる解釈が修飾語が付加詞(Adjunct)になる解釈より優先的である。RAPによると、後位修飾語(Postmodifier)が一番近いヘッドを修飾する解釈が優先的である。

Hobbsはこの二つの原理をさらにMost Restrictive Context PrincipleとAttach Low and Parallel Principleに一般化した[Hobbs90]。

### 2.2 ヒューリスティクスによる優先度

Maruyamaらはヒューリスティクスによる意味的な優先度を計算し、曖昧性の解消を行った[Maruyama88]。Nagaoは構文知識を制約として表現し、意味知識を「嗜好」として使い、それらの知識を動的に組み合わせて曖昧性の解消を行った[Nagao90]。その際、解釈の優先度はヒューリスティクスによって計算される。

ヒューリスティクスによる優先度は理論的な保証が欠けているため、信頼性が低い。

### 2.3 ニューラルネットワークによる優先度

[Apolloni89]ではニューラルネットワークによるPP-Attachmentの曖昧性の解消を行っている。ニューラルネットワークによる優先度は、システムの規模が大きくなると学習時間が多くなるので、現時点では実用的とはいえない。

## 2.4 確率による優先度

Fujisakiらは構文解析によって得られる句構造の出現確率を解釈の優先度とし、曖昧性の解消を行った[Fujisaki89]。句構造の出現確率は、トップダウン解析でその句構造をつくり上げる時に適用される構文規則の適用確率から計算される。SuらはFujisakiらと異なる句構造の出現確率を優先度とし、曖昧性解消を行った[Su88]。句構造の出現確率は、ボトムアップ解析でその句構造をつくり上げる時に適用される構文規則の適用確率から計算される。Hindleらは大規模コーパスから推定した名詞と前置詞、及び動詞と前置詞の共起確率を利用し、PP-Attachmentの曖昧性解消を行った。

この他に、確率的なアプローチで自然言語の解析を行う研究[Church88][Church89][Magerman90]もある。

確率による優先度は、数学的な基礎がしっかりしているし、大規模のシステムにおいても統計によって確率を推定することができるので、もっとも実用的といえる。しかし、優先度に影響を与える構文的、意味的な要因を総合的に考察し、確率的なアプローチで優先度を計算する研究はまだない。

## 2.5 その他

Wilksは、優先的意味論(Preference Semantics)と呼ばれる方法で曖昧性の解消を行った[Wilks85]。具体的には、テンプレートと呼ばれる意味表現における語彙のリンクの数によって優先度を決めている。Tsujiiらは、構文解析の規則と別にヒューリスティクスによる優先規則を用意し、曖昧性の解消を行った[Tsujii88]。

## 3 優先度の計算法

この章では、新しい優先度の計算法を提案する。本研究では、英語の解析を取り上げる。曖昧性解消は構造的な曖昧性の解消と語彙的な曖昧性の解消であるとする。構造的な曖昧性は主にPP-Attachmentによるものであるとする。語彙の曖昧性には同音異義性によるものと語の多義性によるものがあるとする。本研究では構造的な曖昧性の解消と語彙的な曖昧性の解消は同じ枠組みの中で行われるので、以下では特にその両者を区別しないで取り上げる。

本研究では以下のように優先度を計算し、曖昧性を解消する。構文意味解析では自然言語の句構造をつくり上げ、同時に句構造に対応する格フレーム<sup>3</sup>もつくり上げる。句構造とそれに対応する格フレームを解釈とする。複数の解釈が得られた時にそれぞれの解釈の格フレームから意味ファジイと意味確率を計算する。また、それぞれの解釈の句構造から構文確率を計算する。それぞれの解釈の意味ファジイと意味確率と構文確率のベクトルを解釈の優先度とする。優先度のもっとも高い解釈を出力する。

本研究では、格フレームにおけるヘッド、格スロット、修飾語の三者の集まりを格要素と呼ぶことにする。

<sup>3</sup>格フレームは深層の格フレームであってもよいし、表層の格フレームであってもよい。本稿では、表層の格フレームを取り上げる。

## 3.1 意味ファジイ

現在までの多くの自然言語処理システムでは、構文、意味解析は制約(Constraint)の充足によって実現されている。具体的には、構文、意味知識を制約として表現し、その制約を充足する自然言語入力構造を解釈とする。解析の際より多くの制約を取り入れれば、得られる解釈に曖昧性が少なくなる。一般的に構文知識は制約として表現し易い。一方、意味知識は制約として表現しにくい。意味知識を制約として表現する時に常に以下のジレンマに陥る。制約が強すぎて可能な解釈を排除してしまうか、制約が弱すぎて多くの曖昧な解釈を残してしまうかのいずれである。文脈によっては意味的に異常な解釈でも解釈可能であるので、意味解析では意味的に異常な解釈をも残し、すべての解釈の解釈される可能性を計算すべきである。意味知識を解釈の優先順位を決定するための知識とすべきである。

近年、言語学ではプロトタイプ理論[Lakoff87][Taylor89]が注目を集めている。プロトタイプ理論は言語カテゴリに関する理論である。言語カテゴリとは、ある条件を満たす言語的な概念の集合のことである。言語カテゴリは構文カテゴリと意味カテゴリからなる。プロトタイプ理論では以下のようなことを主張する。言語カテゴリを素性ではなく、度合いのついたプロトタイプによって規定すべきである。例えば、「家具」という意味カテゴリを度合いのついたプロトタイプ「椅子」、「机」、「テレビ」等によって規定すべきである。格フレームのスロットも一種の意味カテゴリとみなすことができる。本研究では、以上の考え方を格フレームのスロットに適用する。つまり、格スロットを単語(概念)と単語(概念)のそのスロットに属する度合いによって規定すると考える。このよう知識を意味知識として利用し、解析を行う。

本研究では単語(概念)の格フレームのスロットに属する度合いをファジイとし、構文意味解析の時にファジイを利用し、優先度を計算する。今後このファジイのことを意味ファジイと呼ぶことにする。

なお、ファジイを主観確率によって語ることができ、ファジイを主観確率の近似とみなすことができる[仁木91]。ファジイには取扱い易いというメリットがあるので、本研究ではファジイをつかう。付録で主観確率によって以上の度合いを定義する方法について述べる。

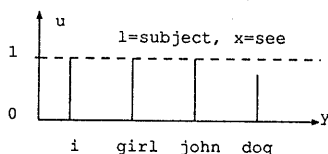


図2: 意味ファジイ

次に意味ファジイを正確に定義する。Aを単語の集合とする。Lを格スロットの集合とする。 $\forall x \in A, \forall l \in L$ 、以下のファジイ集合を定義する。

$$\forall y \in A, 0 \leq u_{x,l}(y) \leq 1 \quad (4)$$

意味ファジイを以下のように解釈する。ファジイ  $u_{x,l}(y)$  はヘッドがxで、格スロットがlである時に、修飾語yが格ス

ロットに属する度合いを表す(図2)。以下では、便宜的に以上の意味ファジイをヘッド、格スロット、修飾語による格要素の意味ファジイという。格要素の意味ファジイはその格要素の意味的な整合性を表す。

自然言語解析システムでは、格要素とその意味ファジイを以下のように記憶する。

<i>see, subject, i</i>	1
<i>see, subject, girl</i>	1
<i>see, subject, john</i>	1
<i>see, subject, dog</i>	0.8

(5)

<i>see, object, i</i>	1
<i>see, object, girl</i>	1
<i>see, object, john</i>	1
<i>see, object, sky</i>	0.8

(6)

<i>see, with, glass</i>	1
<i>see, with, telescope</i>	0.9

(7)

格要素が互いに独立して格フレームを形成すると仮定すれば<sup>4</sup>、格フレームは格要素の論理積によって成り立っていると考えることができる。ファジイ理論では論理式のファジイの最小値を論理式の論理積のファジイとする。よって、本研究では格フレームを構成するすべての格要素の意味ファジイの最小値を格フレームの意味ファジイとする。格フレームの意味ファジイは格フレームの意味的な整合性を表す。格フレームにおいてある格要素の意味ファジイが極端に小さければ、格フレーム全体の意味ファジイも小さくなる。格要素の意味ファジイが平均的に大きければ、格フレーム全体の意味ファジイも大きくなる。これは我々人間の直観と一致する。

式1の文だと、 $u_{see, subj}(i)$ と $u_{see, obj}(girl)$ が充分大きいので、二つの解釈の格フレームの意味ファジイは $u_{see, with}(telescope)$ と $u_{girl, with}(telescope)$ の大きさによる。

$$u_{see, with}(telescope) > u_{girl, with}(telescope) \quad (8)$$

であれば、前者の解釈の格フレームの意味ファジイが大きい。

$$u_{see, with}(telescope) < u_{girl, with}(telescope) \quad (9)$$

であれば、後者の解釈の格フレームの意味ファジイが大きい。格要素の意味ファジイに基づいて、解釈の格フレームの意味ファジイを計算する。

$$see \begin{bmatrix} subject & i & 1 \\ object & girl & 1 \\ with & telescope & 0.9 \end{bmatrix} \quad (10)$$

$$u = 0.9$$

$$see \begin{bmatrix} subject & i & 1 \\ object & girl & 1 \\ with & telescope & 1 \end{bmatrix} \quad (11)$$

$$u = 1$$

<sup>4</sup>当然、すべての格フレームにおいて格要素が互いに独立するわけではない。しかし、多くの場合はそう仮定してもよい。

格フレームの意味ファジイだけを解釈の優先度とすれば、後者の解釈がもっともらしいことになる。意味ファジイによる優先度は意味的な制約の拡張になる。意味ファジイを利用すれば、

Time flies like an arrow (12)

I ate an icecream with a spoon (13)

I ate an icecream with chocolate (14)

I ate an icecream with my friends (15)

のような文に対して、正しい解釈を優先的に出力するだけでなく、意味的な制約で取り扱いにくい意味解析も実現できる。例えば、以下の文はそうである。

I ate an icecream with my finger (16)

I ate an icecream with scraps (17)

I ate an icecream with my cat (18)

### 3.2 意味確率

意味ファジイだけを用いて優先度を計算するのはまだ不十分である。例えば、

He killed a girl with a knife (19)

という文には、「彼はナイフで女の子を殺した」と「彼はナイフをもつ女の子を殺した」の二つの解釈がある。人間は前者の解釈を好む。これはkillとknifeの間につき共起関係があることによると思われる。意味ファジイだけを見ると、前者の解釈の意味ファジイが後者の解釈の意味ファジイより大きいとは限らない。意味ファジイだけでは人間が好むことを表現できない。ある修飾語があるヘッドのある格スロットに属し得る度合いとその修飾語がその格に属する頻度とは別のものである。意味ファジイだけでなく、意味確率を定義する必要もある。

意味確率を以下のように定義する。 $\forall x \in A, \forall l \in L, \forall y \in A$  に対して、意味確率

$$0 \leq P(y|x, l) \leq 1 \quad (20)$$

を定義する。意味確率 $P(y|x, l)$ はヘッドが $x$ で、格スロットが $l$ 、修飾語が $y$ である時の確率である。以下では便宜的に以上の意味確率を格要素の意味確率と呼ぶことにする。

自然言語の解析システムでは、格要素の意味確率を推定し、格要素とその意味確率を以下のように記憶する。

<i>see, subject, i</i>	0.3
<i>see, subject, girl</i>	0.3
<i>see, subject, john</i>	0.3
<i>see, subject, dog</i>	0.1

(21)

<i>see, object, i</i>	0.2
<i>see, object, girl</i>	0.4
<i>see, object, john</i>	0.3
<i>see, object, sky</i>	0.1

(22)

<i>see, with, glass</i>	0.2
<i>see, with, telescope</i>	0.8

(23)

格フレームを構成する格要素の意味確率の積を格フレームの意味確率とする。これは格要素が互いに独立であるという前提に基づく。格フレームの意味確率は格フレームにおける単語の共起頻度を表す。

各格要素の意味確率が大きければ、格フレームの意味確率も大きい。式1の文では、二つの解釈の格フレームの意味確率が  $P(\text{telescope}|\text{see}, \text{with})$  と  $P(\text{telescope}|\text{girl}, \text{with})$  の大きさによる。

$$P(\text{telescope}|\text{see}, \text{with}) > P(\text{telescope}|\text{girl}, \text{with}) \quad (24)$$

であれば、前者の解釈の格フレームの意味確率が大きい。

$$\text{see} \begin{bmatrix} \text{subject } i & 0.3 \\ \text{object } \textit{girl} & 0.4 \\ \text{with } \textit{telescope} & 0.8 \end{bmatrix} \quad (25)$$

$$P = 0.096$$

$$\text{see} \begin{bmatrix} \text{subject } i & 0.3 \\ \text{object } \textit{girl} \left[ \text{with } \textit{telescope} \ 0.1 \right] & 0.4 \end{bmatrix} \quad (26)$$

$$P = 0.012$$

PP-Attachmentの曖昧性解消だけを目的としたら、前置を含む格要素の意味確率だけを解釈の優先度としてもよい。例えば、Hindleらの研究はそうである[Hindle91]。

英語には、名詞に係りやすい前置詞と動詞に係りやすい前置詞、或いは両方に同じくらいに係る前置詞がある。例えば、ofはほとんど名詞にしか係らない。また、withは名詞と動詞に同じくらいに係る。従って、実際の自然言語処理システムでは、二種類の格要素の意味確率を定義してもよい。具体的には、名詞と動詞に同じくらいに係る前置詞を含む格要素の意味確率を  $P(y|x, l)$  と定義する。名詞、或いは動詞に係りやすい前置詞を含む格要素の意味確率を  $P(l|x)$  とする。Hindleらは、 $P(l|x)$  に相当する確率を用いて、PP-Attachmentの曖昧性の解消を行った。

次に、意味ファジイと意味確率の関係について述べる。十分多くのデータから推定できた格要素の意味確率は、格要素の意味ファジイの表すことをも表現する。意味ファジイの小さい格要素の意味確率も小さい。特に、格要素の意味ファジイが0の時、その格要素の意味確率も0になるはずである。従って、意味ファジイを利用せず、意味確率だけで優先度を計算するという考えも原理的には可能である。しかし、意味ファジイの小さい格要素の出現頻度が極めて小さいので、その格要素の意味確率の推定には膨大なデータが必要である。それだけのデータを用意することが現実的に困難である。従って、格要素の意味ファジイと格要素の意味確率を両方利用す

ることが得策である。格要素の意味ファジイは人間が与えるし、さらによく出現する格要素の意味確率はデータから推定する。

### 3.3 構文確率

意味ファジイと意味確率を用いて優先度を計算するのは依然不十分である。例えば、

John phoned a man in Chicago (27)

という文には二つの解釈がある。一つは「ジョンはシカゴにいる人に電話した」で、もう一つは「ジョンはシカゴで人に電話した」である。人間は前者の解釈を好む[Hobbs90]。意味ファジイと意味確率による優先度は人間のこのような優先的な解釈を表現することができない。構文確率を定義する必要もある。

構文確率をいくつかの方法で定義することができる。本研究ではFujisakiらの方法を採用する。具体的には、自然言語を解析する時に得られる句構造の出現確率を構文確率とする。まず、構文規則の確率を定義する。構文規則が  $x \rightarrow y$  である時、条件付き確率  $P(y|x)$  を構文規則の確率とする。この条件付き確率はトップダウン的に自然言語を解析する時に構文規則の適用される確率である。句構造の構文確率はその句構造を導出する際に適用された構文規則の確率の積であるとする。例えば、式1の文の一つの解釈の句構造の構文確率は以下のとおりである。

$$P = P(NP, VP|S) \times P(N|NP) \times P(VP, PP|VP) \times P(V, NP|VP) \times P(DET, N|NP) \times P(NP, P|PP) \times P(DET, N|NP) \times P(i|N) \times P(\textit{see}|V) \times P(a|DET) \times P(\textit{girl}|N) \times P(\textit{with}|P) \times P(a|DET) \times P(\textit{telescope}|N) \times \quad (28)$$

解釈の句構造の構文確率はその句構造の出現頻度を表す。

### 3.4 統合

次に意味ファジイ、意味確率、構文確率を統合して優先度をきめる方法を考える。

本研究では、意味ファジイと意味確率と構文確率をまとめてベクトルとし、このベクトルを優先度とする。最終的に解釈の優先順位をベクトルとしての優先度によって決める。これは以下の考えによる。意味ファジイと意味確率と構文確率は互いに独立したものではない。意味ファジイは意味確率に影響を与える。意味確率は構文確率に影響を与える<sup>5</sup>。従って、意味ファジイと意味確率と構文確率を別々に計算し、最終的にそれらに基づいて優先度を総合的に判断した方がよい。

ベクトルとしての優先度に基づいて以下のように解釈の優先順位を決める。

<sup>5</sup>これは他の研究ではニューラルネットワークによって優先度を計算する理由でもある。

$$\begin{aligned}
& [SeF1, SeP1, SyP1] \geq [SeF2, SeP2, SyP2] \\
& \text{if } SeF1 \geq SeF2 \\
& \text{else if } SeF1 = SeF2, SeP1 \geq SeP2 \\
& \text{else if } SeF1 = SeF2, SeP1 = SeP2, SyP1 \geq SyP2
\end{aligned} \tag{29}$$

但し、 $SeF$  は意味ファジイを、 $SeP$  は意味確率を、 $SyP$  は構文確率を表す(以下同様)。また、ここでいう等しいことはある  $\pm\alpha$  範囲内で等しければよい。このような優先度は人間が構文より意味を重視して優先的な解釈を決めることと対応する。意味ファジイと意味確率は Minimal Attachment Principle を実現していると考えられる。また、構文確率は Right Association Principle を実現していると考えられる。[Hobbs90] によれば、Minimal Attachment Principle が Right Association Principle より優位である。以上で示す解釈の優先順位を決め方はそのことを反映している。

本研究と違って、優先度を以下のように計算することも考えられる。

$$Pref = w_1 \times SeF + w_2 \times SeP + w_3 \times SyP, \tag{30}$$

なお、以上の重み  $w_i$  を経験的に決めるとする。意味ファジイと意味確率が構文確率より重要であることを、例えば、 $w_1 = 5$ 、 $w_2 = 3$ 、 $w_3 = 2$  にすることによって、実現することができる。

以上で示す優先度の計算法は決していびつなものではない。構文ファジイがないのは、構文要素の構文カテゴリに属する度合いをすべて 1 と 0 によって表現できるから<sup>6</sup>である。

### 3.5 学習

以上で示す優先度を計算するためにファジイと確率の初期値を学習によって獲得する必要がある。

格要素の意味ファジイの値を教示学習によって学習することができる。例えば「I see something」が意味的に正しいかどうか(その格要素の意味的な整合性)を人間に質問する。人間がその文の意味的な整合性を 5 段階で評価する。人間が 1 と答えたら、 $u_{see,sub}(i)$  の値を 1 にする。評価値に個人差があるため、複数の人間に質問し、それらの人間の評価値の平均をとって評価値とすることも考えられる。意味ファジイの効率的な学習方式については[李 92]で述べることにする。

格要素の意味確率の値を統計的な手法によって学習する。構造的な曖昧性の解消のためには、表層の単語の共起確率を推定すればよい。語彙的な曖昧性の解消のためには、表層の単語ではなく、語義の共起確率を推定する必要がある。正しく推定するためには人間が予め単語に語義番号を振る必要がある。語彙の共起確率の推定にはコストがかかる。

構文規則の構文確率の値は[Fujisaki89]で示したアルゴリズムによって学習することができる。

### 3.6 メリットと問題点

本研究で提案する優先度の計算法には以下のメリットがある。

<sup>6</sup>[Taylor89]によれば、構文要素の構文カテゴリに属する度合いを 1 と 0 によって表現できない時もある。

1. 本研究の優先度は信頼性が高い。ファジイと確率は人間の不確定問題に対する判断のいくつかの特徴を抽象化したものである。本研究の観察では、人間の自然言語の優先的な解釈がファジイと確率の表現する不確定性問題の特徴をもつ。従って、ファジイと確率による優先度を計算し、解析を行えば、人間と同じ、或いは人間に近い優先的な解釈をすることができる。
2. 本研究の優先度は学習可能である。ヒューリスティクスによる優先度は学習が困難であるが、確率による優先度は簡単に学習できる。学習によって、システムの成績がますますよくなる。また、学習した確率を異なるシステムで利用できることもその優れた点である。
3. システムによっては、意味ファジイ、意味確率、構文確率のどちらかを省略し、残りのものに基づいて優先度を計算してもかまわない。
4. 従来の解析方式との整合性がよい。本研究の優先度は従来の自然言語解析システムに簡単に取り入れることができる。

本研究で提案する優先度には以下の問題点がある。

1. 意味ファジイが主観的なものである。異なる定義者が同じ表現に対して、異なるファジイの値を与えることがある。
2. 意味確率を膨大な量のデータから推定する必要がある。

## 4 おわりに

本稿では自然言語解析における曖昧性解消のための優先度の計算法を提案した。

以上の優先度の計算法をパーザ SAX[松本 86]に取り入れ、自然言語の解析を行った。現段階では主に意味ファジイに基づいて優先度を計算している。意味ファジイだけにに基づく優先度を利用した解析でも、かなりよい結果が得られることがわかった。今後本格的に統計を行い、意味確率と構文確率を推定し、実験をしていく予定である。

本研究は主に以下の点で従来の研究の発展になっている。

1. 意味ファジイを提案した。
2. 意味ファジイ、意味確率、構文確率によって総合的に優先度を計算する方法を提案した。
3. (付録で)主観確率による意味的な優先度について考察を行った。

## A 主観確率と意味ファジイ

プロトタイプ理論によれば、人間は言語カテゴリをプロトタイプ及びプロトタイプのその言語カテゴリに属する度合いによって規定する。例えば、flyの主語をbirdが1の度合いでそのカテゴリに属する、ballが0.4の度合いでそのカテゴリ

りに属する等によって規定する。以上の度合いをどのように定式化すればよいであろう。

本研究で注目したのが主観確率である。主観確率 (Subjective Probability) はいわゆるベーズ流の確率である [Pearl88]。普通の確率が事象の起きる頻度を表すのに対して、主観確率は人間の不確定問題に対する信念の度合い (Degree of Belief) を表す。プロトタイプのカテゴリに属する度合いも人間の信念の度合いなので、主観確率を用いてそのような度合いを表すことができる。以下では、主観確率によってプロトタイプがカテゴリに属する度合いを定式化する。

ベーズ流の確率理論では、主観確率が条件つき確率で、背景知識によってその主観確率が変化すると考える。また、実際の応用では、ベーズの定理を用いて主観確率を計算する。例えば、「a X flies」が正しいことを  $F$  で表す。背景知識を  $K$  で表す。背景知識に基づいて「a X flies」が正しいと判断する主観確率は

$$P(F|K) = \frac{P(K|F)P(F)}{P(K)} \quad (31)$$

である。背景知識  $K$  を考えてみよう。背景知識には二種類のものが考えられる。一つは、概念  $X$  に関する知識である。今  $M$  で「空中で移動すること」を表し、 $W$  で「翼をもつこと」を表す。もう一つの背景知識はその表現が慣用的であるかどうかに関する知識である。ここでは、ある表現が慣用的であることをその表現を聞いたことがあることによって表現する。 $H$  でその表現を聞いたことがあることを表す。人間は異なる背景知識のもとで異なる主観確率で不確定問題に対する判断を下す。例えば、「a bird flies」が正しいことを判断するための背景知識に  $M$ 、 $W$ 、 $H$  があるとすると、つまり、bird が翼をもって空中で移動する、しかも「a bird flies」という表現を聞いたことがあるとすると、それらの背景知識のもとで人間は  $P(F|M, W, H)$  の主観確率で「a bird flies」が正しいと判断する。以下では主観確率  $P(F|M, W, H)$  の値を求める。

$$O(F) = \frac{P(F)}{P(F^c)} = \frac{P(F)}{1 - P(F)} \quad (32)$$

が事前オッズである。

$$O(F|M, W, H) = \frac{P(F|M, W, H)}{P(F^c|M, W, H)} = \frac{P(F|M, W, H)}{1 - P(F|M, W, H)} \quad (33)$$

が事後オッズである。また

$$\lambda(M, W, H|F) = \frac{P(M, W, H|F)}{P(M, W, H|F^c)} \quad (34)$$

が尤度比である。ベーズの定理によれば、

$$O(F|M, W, H) = \lambda(M, W, H|F)O(F) \quad (35)$$

事前オッズ  $O(F)$  が 1 であるとする。つまり、背景知識がまったくない時、事前確率  $P(F)$  と  $P(F^c)$  が共に 0.5 であるとする。また、 $M$  と  $W$  が  $H$  と独立であるとする。よって、

$$\lambda(M, W, H|F) = \lambda(M, W|F)\lambda(H|F) \quad (36)$$

尤度比がわかれば、事後オッズ  $O(F|M, W, H)$  を求めることができる。尤度比を推定するための条件つき確率を以下のように推定できたとする。

$$\begin{array}{|c|c|} \hline P(M, W|F) & 0.6 \\ \hline P(M, W|F^c) & 0.1 \\ \hline \end{array} \quad (37)$$

$$\begin{array}{|c|c|} \hline P(H|F) & 0.8 \\ \hline P(H|F^c) & 0.1 \\ \hline \end{array} \quad (38)$$

ここでは、例えば、 $P(M, W|F) = 0.6$  は人間の「a X flies」という表現が正しければ、0.6 の可能性でその  $X$  が翼をもって空中で移動する」という直観を表す。 $P(H|F) = 0.8$  は人間の「a X flies」という表現が正しければ、0.8 の可能性でその表現を聞いたことがある」という直観を表す。

以上の式に基づいて、

$$O(F|M, W, H) = 6 \times 8 \times 1 = 48 \quad (39)$$

「a bird flies」が正しいこと的主観確率  $P(F|M, W, H)$  は 0.98 である。

一方、「a ball flies」が正しいことを判断するための背景知識に  $M$ 、 $W^c$ 、 $H^c$  があるとすると、つまり、ball が翼なしで空中で移動する。しかも、「a ball flies」という表現を聞いたことがないとする。そうすると、人間はそれらの背景知識のもとで主観確率  $P(F|M, W^c, H^c)$  で「a ball flies」が正しいと判断する。

$$\begin{array}{|c|c|} \hline P(M, W^c|F) & 0.3 \\ \hline P(M, W^c|F^c) & 0.1 \\ \hline \end{array} \quad (40)$$

$$\begin{array}{|c|c|} \hline P(H^c|F) & 0.2 \\ \hline P(H^c|F^c) & 0.9 \\ \hline \end{array} \quad (41)$$

$$O(F|M, W^c, H^c) = 3 \times \frac{2}{9} \times 1 = 0.67 \quad (42)$$

「a ball flies」が正しいことの主観確率  $P(F|M, W^c, H^c)$  は 0.4 である。背景知識が変われば、「a X flies」に対する判断も変わる。

主観確率によって人間の直観と非常に近いプロトタイプのカテゴリに属する度合いを定式化することができた。背景知識がもっと複雑になる場合、背景知識をベーズの推論ネットワークによってモデル化することも考えられる。岩山らはベーズの推論ネットワークを用いて換喩の解析を行った [岩山 92]。

最近、何人かの研究者が主観確率とファジイの関係について研究を行っている。[仁木 91] によれば、ファジイを主観確率の荒い近似と見なすことができる。

主観確率、或いはファジイの値を決めるのに、背景知識のモデルに基づいて計算する方法と、人間が直接それらの値を与える方法が考えられる。以上で示すベーズの定理によって主観確率を計算する例は前者の方法にあたる。実用の立場からみれば前者の方法はそれほど意味があるわけではない。主観確率、或いはファジイの値を人間が直接与えたほうがよい。

また、主観確率を用いるより取扱いやすいファジイを用いた方が便利である。従って、本研究では主観確率ではなく、ファジイによって意味的な優先度を表現している。

## 参考文献

- [Apolloni89] B. Apolloni, G. Mauri, C. Trevisson, P. Valota, A. Zanaboni, *PP-Attachment disambiguation in Natural Language Processing through Neural Networks*, unpublished, 1989
- [Church88] K. Church, *A Stochastic Parts Program and Noun Phrase Parser for Unrestricted Text. The Second Conference on Applied Natural Language Processing*, 1988
- [Church89] K. Church, P. Hanks, *Word Association Norms, Mutual Information, and Lexicography*. *ACL89*, 1989
- [Frazier79] L. Frazier, J. Fordor, *The Sausage Machine: A New Two-Stage Parsing Model*, *Cognition*, 6.191-325, 1979
- [Fujisaki89] T. Fujisaki, F. Jelinek, J. Cocke, E. Black, T. Nishino, *A Probabilistic Parsing Method for Sentence Disambiguation*, *International Parsing Workshop'89*, 1989
- [Hindle91] D. Hindle, Mats Rooth, *Structural Ambiguity and Lexical Relations*. *ACL91*. 1991
- [Hobbs90] J. R. Hobbs, *Two Principles of Parse Preference*, *COLING90*, 1990
- [岩山 92] 岩山真, 江里口善生, 田中穂積 *Bayesian Network* を用いた換喩解析, 1992年度人工知能学会全国大会
- [Lakoff87] G. Lakoff, *Women, Fire, and Dangerous Things, What Categories Reveal about the Mind*. *The University of Chicago Press, Chicago and London*, 1987
- [李 92] 李航, ハイパー・シソーラスとその学習, 情報処理学会自然言語処理研究会, 92-11, 1992
- [Magerman90] D. Magerman, M. Marcus, *Parsing a Natural Language Using Mutual Information Statistics*, *AAAI*, 1990
- [Maruyama88] N. Maruyama, M. morohashi, S. Ueda, E. Sumita, *A Japanese Sentence Analyzer*, *IBM Journal of Research and Development*, Vol. 32, No. 2, 1988
- [丸山 91] 丸山宏, 萩野紫穂, 渡辺日出雄, 確率的形態素解析, 日本ソフトウェア科学会第8会大会, 1991
- [松本 86] 松本裕治, 杉村領一, 論理型言語に基づく構文解析システム *SAX*, *コンピュータソフトウェア*, Vol.3, No.4, 1986
- [Nagao90] K. Nagao, *Dependence Analyzer: A Knowledge-Based Approach to Structural Disambiguation*, *COLING90*, 1990
- [長尾 92] 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, 情報処理学会, 1992
- [仁木 91] 仁木直人, 確率の立場から見たファジィ理論, 日本ファジィ学会誌, Vol.3, No.4, 1991
- [Pearl88] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann, 1988
- [Su88] K. Su, J. Chang, *Semantic and Syntactic Aspects of Score Function*, *COLING88*, 1988
- [辻井 86] 辻井潤一, 文解析方式, 情報処理, Vol. 27, No.8, 1986
- [Tsuji88] J. Tsujii, Y. Muto, *How to get Preferred Readings in Natural Language Analysis*, *COLING88*, 1988
- [Taylor89] J. Taylor, *Linguistic Categorization, Prototypes in Linguistic Theory*. Clarendon Press, Oxford, 1989
- [Wilks85] Y. Wilks, X. Huang, D. Fass, *Syntax, Preference and Right Attachment*, *IJCAI*, 1985

## 謝辞

本研究の機会を与えて下さった NEC C&C 情報研究所情報基礎研究部の中村勝洋部長に感謝いたします。本研究を進めるにあたって、有益なコメントを数多く頂いた安倍直樹主任 (NEC C&C 情報研究所情報基礎研究部) 等、多くの方々に感謝します。