

1E-7 枝分かれするネットワークを用いた格構造の表現と解析

Expression and Analysis of Case Structure with Branched Network

竹内俊行 荒木健治 栃内香次
 Toshiyuki Takeuchi Kenji Araki Koji Tochinai
 北海道大学大学院工学研究科
 Graduate School of Engineering, Hokkaido University

1 はじめに

日本語の意味解析において格構造解析は不可欠であり、格構造を表現するために格助詞は重要な役割を担っている。

従来の意味解析手法 [1, pp.42-94][2, pp.102-157] では文法を基にして解析を行なうために、非文法的な文などの解析が困難である。そこで本稿では、単語とその表現する概念とをノードとするネットワーク [3] を用いて格助詞による格構造を解析する手法を提案し、実験を行ない、考察する。本手法ではネットワークは枝分かれするリンクを持っており、その活性化の状態によって格構造を表現する。このようなネットワークを用いることにより、従来手法では難しい、非文法的な文などに柔軟に対応し、解析が可能である。

2 格構造表現と解析

2.1 枝分かれするネットワーク

本手法では、ネットワークのノードに単語や概念を対応させる。リンクはノード同士の関係付けを表し、これが活性化することでノード間に関係が有することを表現する。そのためにリンクも活性値を持ち、ノードと同様に扱うことができる。例えば、リンクへのリンクを張ることによりノード同士の関係とノードとを直接関係付けることができる。

2.2 格助詞による格構造

格助詞「が」は主に動作主格¹を表すが、対象格も表す。同じように「を」は対象格、場所格、時間格、源泉格などを表す。このように助詞は多くの格を導くが、動詞、名詞、助詞の関係によって、導かれる格を限定することができる。

2.3 格構造表現

本手法で用いるネットワークは言語層と概念層に分かれ、ノードは言語層では単語に、概念層ではある1つの概念にそれぞれ対応している。

システムは言語ノードとその語の意味を表す概念ノードとの間や概念ノード間にリンクを張る。これらリンクの情報は辞書に登録されている。とくに、格助

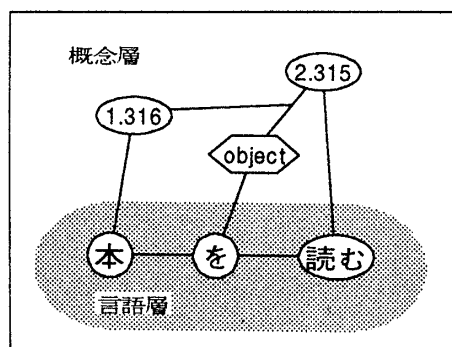


図1 「本を読む」

詞は「格概念」を意味していると考え、これらと概念層の「格概念ノード」との間にリンクを張る。

この格概念ノードと他のノードのリンクが文の格構造を表現する。「本を読む」という文を例にとると図1のようになる。ここで、「本」の概念表現「1.316」は、「を」や「読む」の概念表現「object」「2.315」と直接はリンクせず、「object」と「2.315」を結ぶリンク「{object-2.315}」にリンクしている。²

ネットワークのこのような形状が格構造を直接意味しており、「2.315」の「object」が「1.316」、つまり、「読む」の対象格が「本」であるということになり、「本を読む」の意味を表現している。

2.4 格構造解析

文が与えられると言語層中の該当する言語ノードが活性化する。また、システムにより言語ノード間に直列にリンクが張られ、言語ノードは互いに活性化し、また、それぞれにリンクする概念ノードを活性化させる。概念層では、概念層ノード間で活性化が連鎖的に繰り返される。

一つの言語ノードが2つ以上の意味を持ち、2つ以上の概念ノードへリンクを張っていると、どちらか一方だけが残るようにシステムによりネガティブリンクが張られる。

²ここで、「A」という形式で書かれているものはノードを表し、「{A-B}」は「ノードAとBを結ぶリンク」を表している。また、「1.316」等の数字のノード名は分類語彙表 [4] の分類によるものであり、「object」等は格概念ノードとして本手法で格構造を表現するために導入したノードである。

¹ここで、格は Fillmore の深層格より 11 種類を用いている。[1][2]

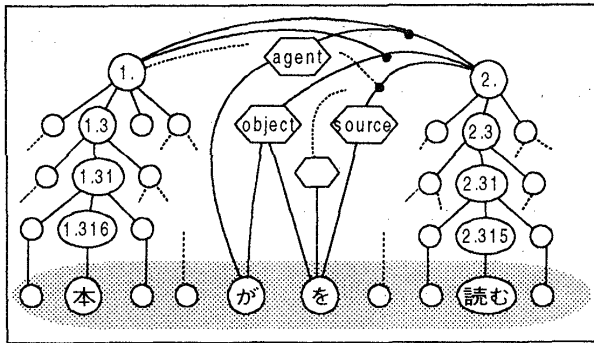


図2 辞書の構成

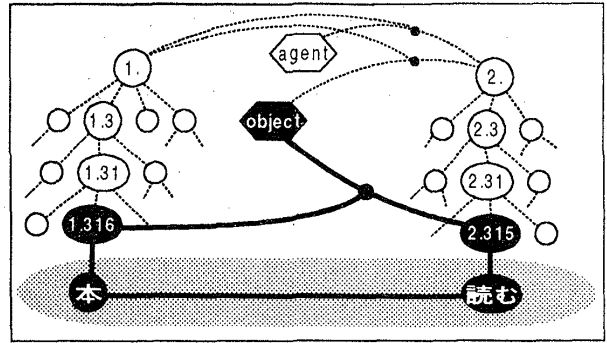


図3 「本読む」

また、活性化したノードとリンクが輪のようになると、そのノードとリンクはより強く活性化されるようになる。そこで活性化値の和を一定にすると、他のノードやリンクの活性化が静まり一部のノードとリンクの活性化値が強く残るように収束する。このネットワークの状態が文の解析結果となる。

3 実験

以上のシステムをワークステーション上に作成し、実際に文の解析を行なった。解析には [5] の文を使用した。

3.1 辞書

辞書はシソーラス [4] を用いて作成し、図2のようになっている。概念層では、'1.' から名詞概念ノード、'2.' から動詞概念ノードが階層状に広がっている。また、それらとは独立して格概念ノードがある。'1.' '2.' とすべての格概念ノードは枝分かれするネットワークによって結ばれている。また、言語層では、各言語ノードからその意味を表す概念ノードへリンクが張られている。また、一文を解析するごとに、その解析結果を新しいリンクとして辞書に追加する。

3.2 解析例

以下では「本を読む」という文の解析結果が辞書に登録された状態での解析結果を記す。

例1: 「本読む」という文を解析するとネットワーク全体は図3のようになる。ここで、「本を読む」の解析結果により '2.315' が 'object' にリンクを張るようになっているために、言語層に「を」などの 'object' を活性化させる格助詞がなくても 'object' が活性化される。このように、不完全な文であっても、それまでの文の解析結果を用いて意味を補うことができる。

例2: 「言葉を話す」という文を解析するとネットワーク全体は図4のようになる。「話す」は '2.' までリンクし、'2.' はすべての格概念ノードとリンクするが、'2.315' が 'object' を、'1.316' が '{object-2.315}' を活性化するため、'object' が他の格概念ノードよりも強く活性化し 'object' が生き残る。このネットワークの状態は、

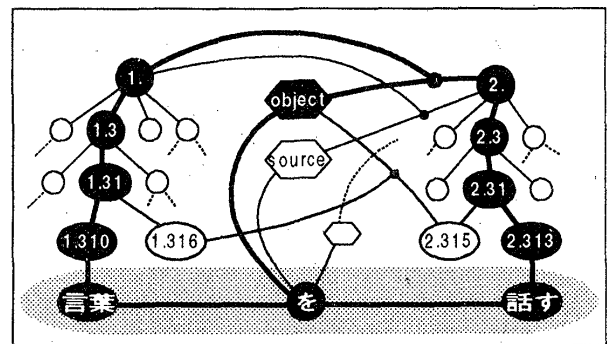


図4 「言葉を話す」

「話す」の対象格が「言葉」であることを表し、正しく解析されている。このように、シソーラスから意味の類似性を用いて文の解析をすることができる。

4 おわりに

本稿では枝分かれするネットワークを用いて、その活性化の状態をそのまま意味の表現とし、また、それを解析する手法を提案した。

本手法では、ネットワークを用いることにより従来の手法では難しい不完全な文の意味の補間などを語の類似性を用いて行なうことができる。しかし、言語層ノード間のリンクの構造や辞書の作成方法などの問題などが残っており、今後の研究課題である。

参考文献

- [1] 石崎俊: “自然言語処理”, 昭晃堂 (1995)
- [2] 田中穂積, 辻井潤一: “自然言語理解”, オーム社 (1988)
- [3] Waltz, D.L. et al.: “Massively Parallel Parsing: A strongly interactive model of natural language interpretation”, *Cognitive Science*, 9, 51-74 (1985)
- [4] 国立国語研究所: “分類語彙表”, 秀英出版 (1964)
- [5] 益岡隆志, 田窪行則: “日本語文法 セルフマスター シリーズ 3 格助詞”, くろしお出版 (1987)