

名詞述語文からの概念学習システム における継承処理

3Q-7

富永 睦 荒木健治 桃内佳雄

北海学園大学

1. はじめに

我々は、自然言語を理解するシステム構築のための1つの基礎として、日本語文の基本的な表現形式の1つである名詞述語文¹⁾を対象とした概念学習システムについての研究を進めてきた^{2, 3)}。本稿では、これまでの評価実験の結果を踏まえた上で、継承処理³⁾の導入によるシステムの改良とその評価実験の結果について報告する。

2. 名詞述語文

名詞述語文とは、主語と述語の対立の中で述語が名詞で形成されている文のことであり、「<名詞句1>は<名詞句2>です。」が基本文型である¹⁾。本稿では、名詞句概念間の基本的な意味関係のうち、下位・上位関係、要素・集合関係、同一関係、対象・属性関係の4つの関係のみを取り扱う。また、名詞句の内部構造として、「~の~」と「~と~」の2つの表現を対象とする。ただし、それらの繰り返し、あるいは埋め込みの表現は対象外とする。

3. 知識表現

知識表現として、S.C.Shapiroらによって提案された意味ネットワークSNePS⁴⁾ (Semantic Network Processing System) に基づく表現を用いる。SNePSでは、ノード（節）が命題、対象、属性、関係などの概念を表し、アーク（弧）がそれらの概念間の構造的、意味的なつながりを表し、このノードとアークにより、人間の心的な世界を表現することができる。そして、外的な世界を表現している語彙と、内的な世界を表現したノードを結ぶために、レックス・アークを用いる。このレックス・アークにより、人間の外的な世界と心的な世界を結びつけて表現することが可能となる。

ある概念からある概念までのアークの連なりを、2つの概念を結ぶパス（経路）と呼ぶ。そして、このパスには直接のパスと冗長なパスの2種類がある。前者は2つの概念をただ1つの、概念間の意味関係を表す概念（関係概念）で結ぶパスであり、後者は2つの概念を2つ以上の関係概念を経由して結ぶパスのことである。

概念区分として、入力された名詞句概念がすでにネットワーク上に存在するか否かにより既知概念、

未知概念と区分する。また、名詞句の内部構造が<は乳類>や<太郎>などのように名詞のみで構成されるものを基準概念、<人間の性格>や<男と女>などのように「~の~」と「~と~」という構造を持つものを複合概念と区分する。したがって、上下に未知概念と既知概念、左右に基準概念と複合概念というように概念空間を4つに区分したネットワークを用いる。本稿において、このようなネットワークを概念関係ネットワークと呼ぶ。

4. システムの概要

名詞述語文から概念および概念間の意味関係を学習するには、一文の情報に曖昧さが多く存在するため、その曖昧さに対処する必要がある。そこで本手法では、学習処理とパス簡約化処理³⁾を融合して行う。学習処理ではヒューリスティクスを用いて意味関係の同定を行うが、ヒューリスティクスによる制約を必要以上に強くすると誤った学習を行う可能性が増大するので、ヒューリスティクスだけでは曖昧さの残る部分を解消するために、知識の有効利用と学習規則を自動的に生成できないだろうかという観点に着目してパス簡約化処理を導入した。また、曖昧さが存在する場合には曖昧さを含んだまま概念関係ネットワークの中に表現し、新たな情報（文）が入力された時点で学習処理およびパス簡約化処理により曖昧さを減少させる方法をとる。

本システムの処理の流れを図1に示す。本手法では、名詞述語文が入力されると、文の構造を解析して、<名詞句1>と<名詞句2>の抽出および名詞句の内部構造を解析する。次の学習処理は、名詞句概念間の意味関係を同定するための処理であり、学習ヒューリスティクスを用いて行う。表1に学習ヒューリスティクスの例を示す。この学習ヒューリスティクスは、概念の区分と概念間の意味関係を推論するため、入力文の構造などを網羅的に考えて生成したものである。学習処理の結果、直接のパスの意味関係が一意に決定した場合にはパス簡約化規則を生成する。また、曖昧さが残り、かつ適用可能なパス簡約化規則が存在する場合には、その規則を適用して意味関係を一意に決定する。このパス簡約化規則とは、入力された文の構造を基にした情報のみならず、周囲の概念との接続状況を考慮することにより知識の有効利用につながる手法として導入した。システムがこの規則を自動的に生成・適用することにより、曖昧さの減少だけでなく、汎用性の拡大にもつながると考える。次に継承処理を行う。本稿における継承処理とは、名詞句の内部構造が「名詞A

Inheritance for Learning System of Concepts from Copular Sentences

Mutsumi TOMINAGA Kenji ARAKI Yoshio MOMOUCHI
Hokkai-Gakuen University

S26-W11, Chuo-ku, Sapporo, 064, JAPAN

の名詞B」を対象とした属性の継承を行うものである。詳細については第5章で述べる。最後の評価学習処理は、ネットワーク上で誤った関係として一意に決定されているものを、正しい関係に修復するための処理である。

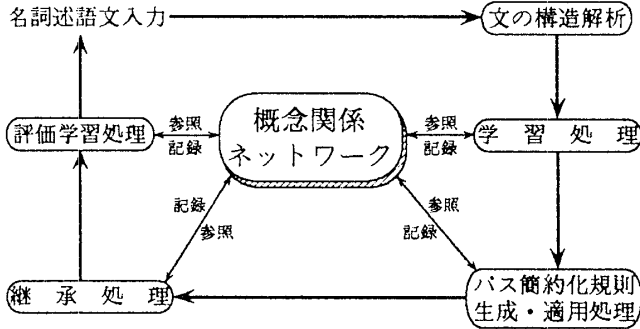


図1 処理の流れ

表1 学習ヒューリスティクスの例

入力文型が「<名詞1>と<名詞2>は<名詞3>です」の場合
 <名詞1>と<名詞2>と<名詞3>の関係→下位・上位関係、同一関係
 <名詞1>と<名詞3>の関係→下位・上位関係、要素・集合関係
 <名詞2>と<名詞3>の関係→下位・上位関係、要素・集合関係

5. 継承処理

本稿における継承処理は、名詞句の内部構造が「名詞Aの名詞B」の場合のみを対象としている。これは、属性を表す<名詞A>をその対象である<名詞B>で制約することにより生成される概念<名詞Aの名詞B>の属性は、<名詞A>の属性のみを表しており、属性の継承を行っても不備が起らないと考えたためである。「人間の性格」という名詞句について考えた場合、<性格>は属性を表しており、その対象を表す<人間>で制約することにより<人間の性格>という概念が生成されている。ここで、<人間の性格>という概念の属性は<人間>という概念の属性のみを表しているため、属性を表す概念<性格>を継承しても不備が起らない。継承処理は、表2に示す継承ヒューリスティクスを用いて行う。

表2 継承ヒューリスティクス

ある概念aを中心に以下の条件を満たす場合、新たな概念εを生成する。そして、概念aと概念εの関係は対象・属性関係、概念εと概念γの関係は下位・上位関係である。
 条件①：概念aと下位・上位関係あるいは要素・集合関係にある概念βが存在する
 条件②：その概念βと対象・属性関係にある概念γが存在する。
 条件③：その概念γと下位・上位関係にある概念δが存在する。
 ただし、概念γの名詞句の構造は「βのδ」である。
 概念εの名詞句の構造は「aのδ」である。

図2は継承処理が行われる状況を示したものである。<花子>と<人間>が要素・集合関係にあり、<人間>と<人間の性格>が対象・属性関係、つまり<人間>の属性に<人間の性格>があり、その<人間の性格>の上位に<性格>がある。このような

状況の下で継承処理が行われ、その動作結果が図3である。属性を表す概念<性格>が継承され、新たな概念<花子の性格>が派生し、<花子>と属性・対象関係、<人間の性格>と下位・上位関係となる。

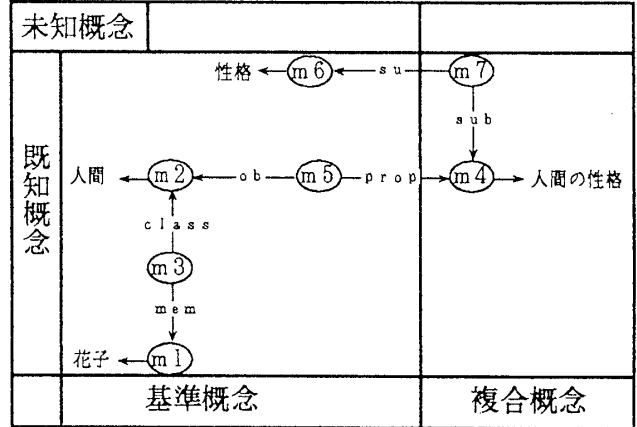


図2 継承処理が起こる世界

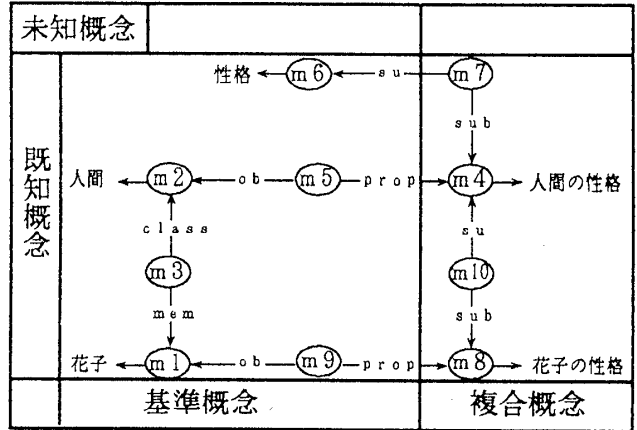


図3 継承処理が行われた世界

6. 評価実験

本手法に基づく実験システムをWS上に作成し、名詞述語文を50文入力して、評価実験を行った。評価実験の結果、2文において継承処理が行われ、いずれも正しい関係で他の概念と結びついている。

7. おわりに

評価実験を行った結果、継承処理が良好に行われていることが確認された。

今後の課題として、名詞句の内部構造の拡張、処理可能な意味関係の増加、多義語の処理などがある。
 *参考文献

- 1) 高橋太郎：名詞述語文における主語と述語の意味的な関係、日本語学、Vol.3, No.12, pp.18-39, 明治書院、東京(1984)。
- 2) 荒木健治、桃内佳雄：名詞述語文における意味概念の学習、情処研報Vol.90, No77, pp.1-8(1990)。
- 3) 佐々木淳一、冨永睦、荒木健治、桃内佳雄：名詞述語文からの概念学習システム、情処研報Vol.93, No1, pp.39-46(1993)。
- 4) S.C.Shapiro: "The SNePS Semantic Network Processing System," in N.V.Findler(ed.), Associative Networks, Academic Press, New York, pp.179-203(1979)。