

換喩を考慮した動詞と格要素の
共起性の判断

5F-7

藤木武博 平井誠 北橋忠宏

大阪大学産業科学研究所

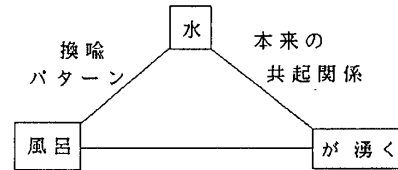
1. はじめに

自然言語解析システムにおける語彙知識の基本的な役割は、語彙の持つ意味的な制約をなんらかの形式で記述して主部と述部、動詞と格要素などの概念間の共起関係を決定して、文の意味的な妥当性を判断することである。しかし、日常的には本来の共起関係を逸脱しながら、非文とはならない、特殊な意味構造を持つ言語現象が使われることがある。その代表的な例が比喩と換喩である。

本論文では、動詞と格要素間に成立する換喩について考察し、筆者らが開発している日本語解析システムMARIONに新たに換喩解析の枠を加えることによって、動詞と格要素の受理範囲を拡張し、人間の主観により近い解析を行う事を検討する。

2. 換喩の構造

比喩は対象となっている事物を別の事物の特性を利用してたとえる表現であり、その解析には文脈的、状況的、語用論的、そしてさらには常識的な情報を統合的に処理する必要がある。そのため計算機上の限られた知識範囲で解析することは非常にむずかしい問題を多く含んでいる。一方、換喩はある対象となっている事物を別の事物で換言する言語現象であり、特に日常会話で頻繁にみられる一種の情報圧縮と考えることができる。動詞と格要素の関係に注目すると、その基本的な構造は「風呂が湧く」のように共起性を逸脱した接続関係(風呂と湧くは本来は共起しない)が圧縮されている情報(ここでは水)を介して意味整合を保っている構造をしている(図2)。換喩の特徴は表層に現れている格要素(以下、仮格要素と呼ぶ)と圧縮されている概念との関係がある程度限定されていることである。そのため、限定されたパターンを基に適切に圧縮されている概念の予測を行うことができれば比較的簡単に解析を行うことができる。



仮格要素

図1 「風呂が湧く」の構造

3. 換喩のパターン

換喩のパターンは山梨^[1]によって指摘されているが、実際的な根拠を得るために換喩例の収集を行った。換喩は特に日常的な状況で現れる言語現象であるので、日常会話を中心にサンプル抽出を行うのが適当と考え、テレビドラマのシナリオを対象に動詞と格要素の関係から換喩とみられる文例を収集した。およそ、150件の文例に対して整理を行った結果、次の五つの代表的なパターンがあることがわかった。[A←B]の表記はAが仮格要素、Bがそれに対応する圧縮情報の概念を表している。また{}は抽出された文例数を表す。

(1) 提供体←被提供物 {26}

例) レコードを聴く

(2) 構造物体←構成部品 {32}

例) 水道をひねる

(3) 被支持物←存在空間提供体 {41}

例) コーヒーを持って来る

(4) 存在空間提供体←被支持物 {15}

例) 部屋を片付ける

(5) 属性値による換喩 {9}

例) 400cc に乗っている。

4. MARIONにおける換喩解析

MARIONでは動詞と格要素との共起関係は名詞ソーラスを用いて個々の動詞について意味規定を行うことにより記述している^[2]。つまり、格要素が意味規定を満足するなら意味的に整合性のある表現として受理する。換喩の解析は格要素が動詞の意味規定を満たさなかった場合、換喩パターンに基づいて圧縮されている概念を予測して、その概念が動詞辞書の意味規定を満足するかどうかを評価することによって行う。満足しなかった場合は非文とする。この際、圧縮概念を予測するための情報源となるのが名詞の意味情報である。現在、MARIONにおいて名詞の意味情報は、独立した視点に基づく概念階層木と概念間の関係を表す意味ネットワークの2つの記述形式を併用している。この制約のもとで換喩を解析する方法を検討した。

(1)のパターンについては必要な情報は名詞の機能に関する情報であり、これはMARIONにおいてはPROVリンクで表現されている。このリンクはある概念が一般的に何を提供するかという関係を表す。仮格要素がPROVリンクによって他の概念と接続している場合には、その概念と処理中の動詞との共起関係を調べる。もし、PROVリンクがない場合は、仮格要素の上位概念について同様の処理をおこなう。この処理を上位概念が現れなくなるまで再帰的に適用する。(2)のパターンについては必要な情報は名詞の構造に関する情報であり、これはMARIONにおいてはHASリンクで表現されている。このリンクは、一般的な「全体-部分」関係を表し、構造体に対してはその構成部品を対応させているので(1)のパターンと同様の処理を行い共起性を調べる。但し、上位概念への再起的適用は行わない。(3)のパターンについては必要な情報は(1)と同様に名詞の機能記述であるが、MARIONの意味リンクは一方的であり、PROVリンクは提供する概念から提供される概念への情報獲得のみが可能なので、リンクを辿ることによっては圧縮概念の予測を行うことができない。しかし、文例中の動詞や名詞の重複度を考慮すると有効な文例数は少ないことや、被提供物がほとんど液体に限

られているため仮格要素が液体の場合には常に容器を圧縮概念として補うことにより共起性の評価を行うものとする。(4)はそのほとんどが2次元空間を提供するパターンの換喩であり、この場合2次元空間によってその存在を支持されている具体物が状況や常識に依存することが多いため、名詞辞書から圧縮情報の予測をすることはむずかしいので常識的な制約のみ動詞辞書に記述する。(5)のパターンは慣用的な性格が強いパターンであるので(4)のパターンと同様の理由で動詞辞書に記述する。これらの処理過程を図2に示す。

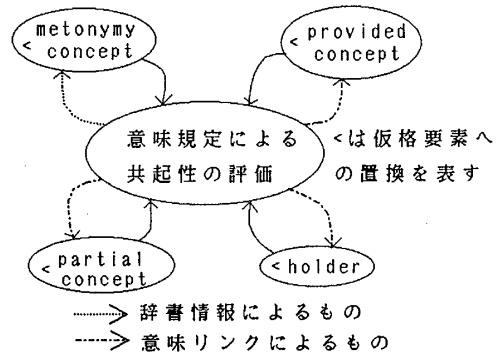
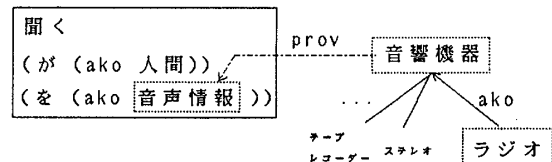


図2

5. 処理例

「ラジオを聞く」は次のようなリンクを辿ることによって共起性を判定される。



6. まとめ

換喩を考慮した動詞と名詞の格要素の判定について考察した。動詞と名詞の共起関係を精密に行うためには、今後は比喩的な接続関係や、意味範囲の広い動詞の意味規定などを検討していく必要があると思われる。

7. 謝辞

本論文において有益な御提言を頂いた北橋研究室の土井敦子技官に感謝致します。

8. 参考文献

- [1] 山梨正明:「比喩と理解」東京大学出版会1988
- [2] 平井誠, 北橋忠宏:「語の意味情報の共通性と個別性」(文法知識と意味知識の管理シンポジウム論文集)電子通信学会, 1989