

スperlベルの象徴解釈モデルに基づく隠喩の検出†

土井 晃 一^{††} 田中英彦^{†††}

自然言語処理に止まらず、自然言語理解を計算機上で行うためには、現在談話理解として盛んに研究されている象徴と指示関係の理解が必要である。そのほかにも、単語のニュアンスの問題を含んだ単語の使い方、言内の意味を越えた言外の意味の解析、発話を取り巻く環境を表すその場の状況理解、文脈理解、話者の認識の仕方の解析等の問題点がある。これらを扱うアプローチの一つとして、隠喩理解がある。我々は、隠喩理解を行う第一段階として隠喩の検出を計算機によって機械的に行う。隠喩の検出には、スperlベルの象徴解釈モデルを用いる。隠喩検出の基準としては、明らかに偽な場合、不条理な場合、有意性を欠く場合の三つに分けて取り扱う。実装システムでは観念的なことを扱うために意味ネットワークを用い、実在的なことを扱うために状況意味論を用いた。また計算機上の実装は Prolog で行った。本論文では、計算機上での隠喩検出に必要なデータ構造と、その検出の仕方について述べる。さらにその検出の範囲についても言及する。

1. はじめに

自然言語処理に止まらず、自然言語理解を計算機上で行うためには、現在談話理解として盛んに研究されている象徴と指示関係の理解が必要である。そのほかにも、単語のニュアンスの問題を含んだ単語の使い方、言内の意味を越えた言外の意味の解析、発話を取り巻く環境を表すその場の状況理解、文脈理解、話者の認識の仕方の解析等の問題点がある。これらを扱うアプローチの一つとして、隠喩理解がある。隠喩とは単語の意味が文脈によって変わる作用のことをいい、字義どおりでは意味が通じないものを指す。

我々は、隠喩理解を行う第一段階として隠喩の検出を行う。隠喩の検出には、従来のグライスの「会話の作法」ではなく、スperlベルの象徴解釈モデルを用いる。隠喩検出の基準としては、明らかに偽な場合、不条理な場合、有意性を欠く場合の三つが挙げられる。

実装に用いたモデルは観念的なことを扱うために意味ネットワークを用い、実在的なことを扱うために状況意味論を用いた。また計算機上の実装は Prolog で行った。

2. 隠喩理解

2.1 隠喩理解の意義・目的

言内の意味だけではなく、言外の意味を発見しない

と会話の理解はできない”。

ここでいう文脈理解とは、外の世界の状況とは無関係に、前後の文章あるいは全体の文章によって文の解釈が変わる場合のことをいう。文脈理解の中には、話し手と聞き手の心の状態、共有知識、先行する発話などが入ってくる。文脈理解も自然言語理解には不可欠である。

一方、その場の状況理解とは、文章の中の世界だけではなく外の世界とのつながりを持つことの意味を指し、話題と直接関係ないことが発話されたときそれが話題とどう関係するのかを調べることである。話者を取り巻く世界に関する情報はすべて状況理解と関係する。世界に関する記述は、特別の指摘のない限り、話し手の発話をそのまま信用して取り込むことにする。

新しい状況に対する話者の認識の仕方の解析とは、現実の世界を話者がどのように理解したかを示すという意味である。話者が未知の対象について陳述しようとする場合、既知の表現では表現しきれないから、その話者に固有の表現が発話されることになる。そうするとその対象に対する話者の認識の仕方が表れることになる。このように隠喩理解を行うことにより、人間の心理的な側面に関する研究を行うことができるようになる。

2.2 隠喩理解の応用

隠喩理解を行うことにより、既存単語の新しい意味、新単語の意味の解析が可能になる。例えば「プログラムが走る」という例を挙げると「走る」という単語に物理的に物体が走るという意味が従来あったが、これに「実行される」という新しい意味が付け加えられたことになる。つまり隠喩理解は新知識獲得の手法でもある。

† Metaphor Detection Based on Sperber's Symbol Analysis Model by KOUICHI DOI (Information Engineering Course, Division of Engineering, Graduate School, University of Tokyo) and HIDEHIKO TANAKA (Department of Electrical Engineering, Faculty of Engineering, University of Tokyo).

†† 東京大学工学系研究科情報工学専攻

††† 東京大学工学部電気工学科

また単語のニュアンスの問題を扱うことを目的として機械翻訳にも活用できる。新聞、雑誌など毎日多量に生産される文献は人間の手を介して翻訳するには量的に無理があるので、翻訳の質向上に機械翻訳の出番となるが、各国語の文化的背景を持った隠喩は頻繁に出現するものであり、また新しい隠喩は次々に誕生し、単語の意味もどんどん変わっていく。このような状況下では隠喩理解は不可欠の要素になる。

また現在の研究に役立つ点としては、言語処理の分野でソーラスを作る際に単語の意味を確定するのに必要となることが挙げられる。

3. 隠喩検出

本論文では、隠喩理解を隠喩検出と隠喩理解に分け、その中で隠喩検出について述べる。その際隠喩だけを検出の対象とするよりも換喩、提喩も含めた、コノテーション(判示的意味作用)を検出の対象とした方が応用上望ましいので、それらをも検出の対象とする。

3.1 隠喩検出

隠喩の検出の原理についてはグライスの提唱した「会話の作法」を用いることが考えられる。以下メタファの記号論から引用すると、これは会話が円滑に進むべく話者が守る作法を定義したもので、初めてその原則を明文化した例である。グライスの会話の作法をカントの判断の分類表に倣って以下に挙げる。

量

- (1) (情報交換の現在の目的にとって) 必要なだけの情報を、君の会話への寄与が持つようにせよ。
- (2) 必要以上の情報を君の寄与が持たぬようにせよ。

質

- (1) 君が偽だと信じることを言うな。
- (2) その十全な証拠が君にないようなことを言うな。

関係

有意的であれ。

様態(この作法は一口に言う「明確に表現せよ」という主旨だが、グライスはこれを次のように細分して示している。)

- (1) 表現の曖昧さを避けよ。
- (2) 両義性を避けよ。
- (3) 簡潔であれ(不必要な冗長を避けよ)。
- (4) きちんと順序よく言え。

この「会話の作法」を用いてグライスは隠喩の検出が行えるとしている。

さらにグライスによると、隠喩の必要十分条件は以下ようになる。

- (1) 隠喩は、すべて、厳密下位範疇化規則に違反しない
- (2) すべての隠喩には、共起場面が欠如している
- (3) すべての隠喩には、アイコン(対象のイメージがはっきりと心に浮かぶこと)³⁾ 的要素が含まれている

しかしグライスの会話の作法を使つと意味論の範囲で隠喩検出をしなければならなくなる。その結果、語の本来の意味と文彩された意味の対立を持ち込むことにより語の意味に曖昧さが生じないが、検索する範囲が定まらないのではないか、さらに文字どおり真であってさらに隠喩的な意味を持つ文はどうするのか等の疑問が提出されている。

3.2 他の関連研究

また本論文で用いる意味ネットワーク⁴⁾、状況意味論⁵⁾については諸文献を参照されたい。またすでに隠喩理解に関してなされた研究もある²⁾。そこではグライスの「会話の作法」を用いて共起場面の欠如の起きたときに隠喩であることを見だし、Prologのユニフィケーションプロセスや、コネクショニストモデルを用いて隠喩の理解を行っている。

しかし現実の発話文に近い形から隠喩を検出し、隠喩の一般的な理解を目指す研究はまだなされていない。

3.3 スペルベルの象徴解釈モデル

スペルベルによると、人間が象徴を解釈する際の構成は図1のように知覚装置、理性装置、象徴装置、記憶の四つからなる。

まず知覚装置とは外的刺激により供給される情報(スペルベルの場合は情報は言語とは限らず、話者の表情、身振りなども入るが)を入力として受容し、すでに経験された事象の知識の総体の中からいま入ってきた情報と一致したものを命題の形でこれを同一指定

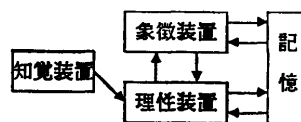


図1 スペルベルの象徴の解釈モデル
Fig. 1 Sperber's symbol analysis model.

して、さらにこれを出力として出す一連の操作の総体をいう。理解装置とは、命題を入力として受け取り、入力（および記憶中で使用可能な前提と）から「論理的に導出された」他の命題を出力として出す一連の総体をいう。最後に象徴装置とは、命題を入力として受け取り、入力より「呼び起こされた」他の命題を出力として出す一連の操作の総体をいう。象徴装置は新たな入力が入るまで無限に象徴の解釈を続ける。そして解釈の結果を理性装置へ送り返す。図1で、実線の矢印がデータの入力を示し、破線のそれがフィードバック回路を表す³⁾。

スペルベル・ウイルソンは意味論を包括する修辞論の立場からグライスの会話の作法を含む、有意性公理を提案した。スペルベル・ウイルソンによると発話された文が

- (1) 明らかに偽である
- (2) 不条理である
- (3) 有意性を欠く

とき呼び起こしが生じ、言外の意味の探索が行われると主張した³⁾。本論文ではこの有意性公理を用いて隠喩の検出を行う⁴⁾。

(1), (2)はグライスの会話の作用の質、様態に（虚偽の発話をしない、明確な表現をするという点から）(3)は量、関係に相当する（無関係なことは言わないという点で）³⁾。

3.4 スペルベルのモデルの利用

本論文では前節で述べたスペルベルの象徴解釈のモデルを利用して隠喩の検出を行う。このモデルは象徴一般を検出し理解するものであるが、ここでは隠喩の検出を行うので、入力は自然言語に限ることとする。

さらにここでは隠喩の検出、理解を目的としているので、知覚装置は一応構文解析が終わったものを出力するものと仮定し、その具体的な構造は扱わない。知覚装置からの出力としては、構文解析された結果として、状況意味論⁵⁾の situation type, state of affair, event type のいずれかをとることとする。

スペルベルによると発話に対して何かおかしいと感じるのが「呼び起こし」、それに対して関連する命題を検索に行くのが「焦点合わせ」である。前者に当たるのを本論文の検出ルーチンとし、後者に当たるのを理解ルーチンとする³⁾。

我々は理性装置の行う範囲の一部を隠喩検出ルーチンと考え、象徴装置の行う範囲の一部を隠喩理解ルーチンと考える。前述のように、本論文の目的はこの検

出ルーチンの構成にあるので、文彩された意味の入力が既存知識に直接干渉しないように記憶は静的な既存知識と動的な入力を保持することが必要であるから、二つに分け、さらに静的な既存知識を意味論の範囲（命題の真偽値を問題とする範囲）³⁾ で扱う部分を意味ネットワークで表し、修辞論の範囲（象徴一般を問題とする範囲）³⁾ で扱う単語のニュアンスを扱う部分を状況意味論で記述することにする。動的な知識（文脈、状況）は、作業領域に保存する。

比喩は、直喩（例えば「りんごのような頬」）、隠喩（「人間は狼である」）、諷喩（ほんとに伝えたいことをそっくり他に移行しかえるやり方）、活喩（無生物を生命のあるものとして扱う比喩）、提喩（全体を部分に、部分を全体で表す比喩）、換喩（概念を属性で、属性を概念で表す比喩）、引喩（他人の言葉を引用するやり方）、張喩（事実を誇張するやり方）、声喩（擬態語、擬声語を使う方法）、字喩（「いとしいとしという心」）、詞喩（掛け言葉など）、類喩（「川」「流れ」「水」「浮く」等という言葉を本文とは無関係に使い、別のイメージを持たせる方法）の十二種に大分される⁸⁾。このうち隠喩と関連の深いのは、隠喩、諷喩、提喩、換喩、引喩の五つである。これらを検出するには前述の呼び起こしの条件を用いればよい。

4. 隠喩検出システムの全体構成

4.1 システムの内部構成

隠喩検出システムは図2に示すように、

- (1) 既存の単語についての知識を保持する静的な意味ネットワーク部
 - (2) 状況、文脈を保存する動的な作業領域
 - (3) 実際に検出を行う検出ルーチン
- とからなる。この(1)はさらに意味ネットワーク部と状況意味論で記述された構造部からなる。

状況の移り変わりは course of event として順次入力されることになる。

また現実のプログラム（意味ネットワークと検出

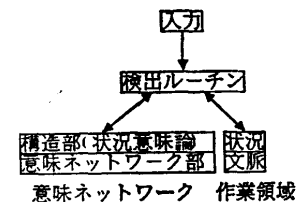


図2 システム構成
Fig. 2 System construction.

ルーチン)は Prolog で記述してある。

4.2 入力

検出ルーチンに入る入力形としては次のような形を仮定する。状況の識別子としては“s”，時空位置として“1”，次に対象の n 項関係が並び，最後にその状況の真理値が記述してある。状況間の関係は，一つの状況が入力され，受理されると次の状況を見に行くという形をとっている。

(1) (名詞) が (名詞) である

例えば「人間は狼である」という発話は

in s : at 1 : is, man, wolf ; yes

という形を仮定する。

(2) (名詞) が (用言) する

例えば「貿易摩擦」という発話は

in s : at 1 : rub, trade ; yes

という形を仮定する。

(3) 目的語を持つもの

例えば「時間を浪費する」という発話は

in s : at 1 : spend, man, time ; yes

という形を仮定する。

(4) 所有形容詞を持つもの

例えば「彼の母親は天使だ」という発話は

in s1 : at 11 : have, B, A ; yes

in s2 : at 13 : is, A, angel ; yes

in s3 : at 14 : is, A, mother ; yes

という形を仮定する。Bが「彼」に，Aが「母親」に当てはまる。Bが人間であるということは入力の際に現実の状況で結びつけられ，無矛盾のチェックが行われる。

(5) 指示形容詞を持つもの

例えば「あの男は狐だ」という発話は

in s1 : at 11 : that, A ; yes

in s2 : at 12 : is, A, fox ; yes

in s3 : at 13 : is, A, male ; yes

という形を仮定する。ここで male は日本語の「男」に当たる。

4.3 意味ネットワークの意味ネットワーク部

意味ネットワーク部は，単語のもっとも基本的な関係である。単語の包含，属性，対立関係だけを表す概念的なネットワーク部と，単語の使用法を詳細に記述した実在的な構造部とからなる。木をすべて探索すれば二つのものの関係は分かるから，対立リンクは不要とも言えるが，隠喩理解では対立概念の認識が不可欠であるため，従来の意味ネットワークにはない対立関

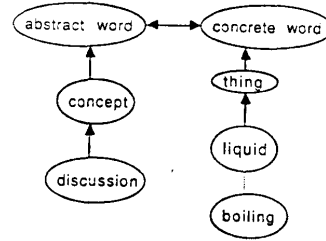


図3 意味ネットワークの例
Fig. 3 Example of semantic network.

係を表すリンクが張ってある。

発話と意味ネットワーク上のノードを区別するために便宜上発話を日本語で，意味ネットワーク上のノードを英語で表してあるが隠喩理解の対象は本論文ではあくまで日本語である。

図3は意味ネットワークから「議論は沸騰する」に関係する部分を切り出してきたものである。丸で囲まれた中の単語が概念のノードに当たり，片方向矢印は包含リンクを表し，矢印の付け根が下位概念，矢印の先が上位概念を表す。双方向矢印は対立リンクを表す。点線は属性リンクを表す。

現実の実装はノードを Prolog のアトムで，リンクをクローズで記述してある。例えば「議論は沸騰する」という例であれば，

be (discussion, concept).

be (concept, a-word).

be (liquid, thing).

be (thing, c-word).

attr (liquid, boiling).

のようになっている。アトムである liquid がユニファイの際 boiling から参照され，意味ネットワークを上位概念にたどることになる。対立関係は対応する包含リンクがないことにより，間接的に表されている。

この意味ネットワークによって隠喩の検出，理解を行うので意味ネットワークの作り方が問題となる。意味ネットワークの作り方については，階層的意味素性体系についての研究⁹⁾，比喩の機構・伝達効果の研究¹⁰⁾などを参考にした。

4.4 意味ネットワークの構造部

構造部は単語に一対一対応し，その単語の使い方を記述してある。実体は単語の使い方を記述したものを要素に持つ situation type の集合である。

意味ネットワーク部の構造部は三つの部分からなる。

- (1) 現象の存在
- (2) 現象の認識
- (3) 発話

現象の存在とは、この現実世界に実際に起きる現象を抽象的な状況として記述してある。先の例を図4に示す。まず発話に先立ち「Aが酒をたくさん飲んだ」と、その後「Aが暴れた」ことを記述してある。次の現象の認識とは、発話に先立って話者が現象の存在を認識したことを記述してある。最後の発話では話者が現実に行う発話を記述してある。

実際には、Prolog で記述されている。例えばだれかがたくさん酒を飲んだという状況は、状況意味論の記述で、

```
in s1: at 11: drink, A, much; yes
```

と記述されるが、これが Prolog のクローズの形で、

```
do (L 1, drink, A, much).
```

と記述されている。本来この状況は抽象的状況であるのでそのことを示す識別子とこの状況に対する識別子が必要であるが、現在の段階では不要なので明示されていない。

また諺に関する情報も後述するように構造部に保存されている。

4.5 作業領域

作業領域は検出ルーチンを経由して入ってきた入力を保存しておく部分で後に隠喩理解の際にもう一度参照することになる。

実際にはこれも Prolog のクローズの形で保存され、例えば juwa が酒をたくさん飲んだという状況は、状況意味論の記述で、

```
in s1: at 11: drink, juwa, much; yes
```

となるが、これは

```
do (real, L 1, drink, A, much).
```

という形で保存されている。ここで *real* は現実的状況であることを表している。

構造部は状況意味論でいうところの抽象的状況が記

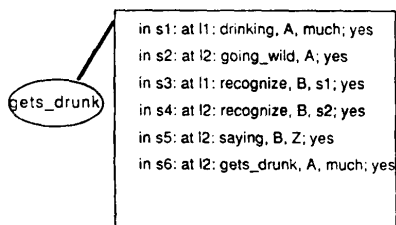


図4 構造部の例

Fig. 4 Example of structure part.

述されているところである。隠喩検出ルーチンはこの抽象的状況に照らし合わせて、実在的状況を作業領域に書き込む。

作業領域に保存されるこれらの状況は、隠喩検出の際に抽象的状況と照らし合わせて実在的状況が存在するかどうかを確認するのに使われるほか、隠喩理解の際も実在的状況の存在という形で使われることになる。

4.6 検出ルーチン

検出ルーチンは入力が入ってくると作業領域内の命題を参照しながら、意味ネットワークを検索し、第5章で述べる検出のアルゴリズムを実行する。

5. 隠喩の検出

5.1 検出の原理

検出の原理はグライスの会話の作法を改良したスペルベル・ウイルソンの「呼び起こし」の原理に基づいて実行される。

スペルベル・ウイルソンによると、聞き手は話し手の発話が明らかに偽であるとき、不条理なとき、有意性を欠いた発言をしているとき「呼び起こし」が生じ、再解釈を始めると分析している。我々は従来のグライスの会話の作法を用いずこの「呼び起こし」原理に基づき、計算機上で隠喩の検出を行う。

全体の流れとしては上の三つの場合を順次調べることになる。検査する順番は、命題レベルのチェック、文脈レベルのチェックその一（現実の状況に反する場合）、文脈レベルのチェックその二（現実の状況に反しない場合）となる。この順番は人間が検査する順番を模倣している。上のどれかに当てはまると隠喩の可能性のあるものということになり、どれにも当てはまらない場合には、このシステムにとっては新知識ということになる。

発話の中の情報だけで真偽が決まる場合を分析的、発話の外、つまり対象世界によって真偽が変わる場合を総合的であるという言い方を使うと、明らかに偽である場合とは分析的で発話の真理値が偽の場合を指し、不条理な場合とは総合的で偽の場合のことであり、有意性を欠く場合とは常に真の場合に当てはまる。

5.2 明らかに偽である場合

まず明らかに偽であるかどうかは（主に動詞、形容動詞、形容詞、名詞などからなる場合）、意味ネットワーク部のネットワーク部で包含関係のチェックと、主語／述語、修飾語／被修飾語関係の間に属性関係が

あるかどうかのチェックを行うことにより判定する。

「(名詞)は(名詞)である」という形の隠喩が包含関係のチェックで検出される場合であるが(「人間は狼である」が例として挙げられる), この場合意味ネットワーク上で包含関係にない, 包含関係が逆になっていることから隠喩であることが分かる(当然話者が虚偽の発言をしていることもあるが, それは隠喩理解の際に動的に決定される)。

「(名詞)が(用言)する」という形の隠喩は主語/述語, 修飾語/被修飾語関係の間に属性関係があるかどうかによって検出される場合であるが(「議論が沸騰する」が例として挙げられる), この場合, 主語あるいは被修飾語の属性に述語あるいは修飾語があるかどうかで隠喩かどうかの判定をする。

5.3 不条理な場合

次に明らかに偽ではない場合には不条理であるかどうかを(主に副詞の場合)意味ネットワークの中の構造部を検索し, 発話に当たる状況が存在するかどうかを調べることで判定する。

これは発話の内容が現実の状況に一致しないときであるが, 例としては古来の分類法によると隠喩というより皮肉の例になるが, 酒をたくさん飲んだため暴れている人を見て「彼は少し酔っている」と発話したとすると, これは言内の意味としては状況にそぐわないわけだが普通の発話としてはよくみられる例である。

酔っているというノードに対応して単語の使い方を記述した構造部の中を検索することになる。構造部の中はいろいろな状況に合わせて, 状況意味論の situation type で記述しており, これが入力状況と一致するかどうかで言外の意味が存在するかどうかを判定する。この場合隠喩的な単語の使われ方も意味ネットワークの中に組み込まれていることになる。

5.4 有意性を欠く場合

最後に明らかに偽な発話でもなく, 不条理でもない場合には, 有意性を欠いているかどうかを判定する。

これには二つの場合が挙げられる。

- (1) 常識のレベルを定めるもの
- (2) 諺に関するもの

常識のレベルとは, その発話の内容が普通の会話の知識内で処理できるかどうかのレベルのことであり, 実装は意味ネットワークの中にすでにそのリンクが存在するかどうかで判定する。ただし専門用語が現れたときには, その単語に関する説明, 確認とみなすことにする。専門用語でも一度発話された知識が二度以上発

話されたときには, 可能性があるものとみなされる。専門用語でない場合でも前提としての発話がなされることがあるが, これが隠喩的なものであるかどうかは, 隠喩理解ルーチンが判定する。

分析的, 総合的という言葉を用いると, (1)は文脈に依存するので総合的, (2)は発話内の情報だけで検索できるので分析的であると言える。

これらは二重に真である場合であり, 文脈とは直接関係のない発話がなされたときである。例えば, (1)の例としては「人間は動物である」, (2)の例としては「隣の花は赤い」などが例として挙げられる。「人間は動物である」という発話では, まず言内の意味が真である。人間が動物であるのは当然であり, 話者はそのことを聞き手に知識として伝えようとしているのではない。言外の意味を主に伝えようとしている。次に「隣の花は赤い」という発話では, 例えば, 隣家の新車の話をしているとき突然この発話がなされたとする。本当に隣の花壇の花は赤いかもしれない。しかし話者は隣の花壇の花が赤い事実を指摘したかったのではなく, 隣家のものはよく見えるということを指摘したかったのだ。この例題を検出するには, 構造部におおのの状況から連想される単語を記述しておく, 諺とのパターンマッチを図ることにより解決される。これについてのより詳しい考察は今後の課題である。

この諺に関するものの中には, 有名な句などの引用も含まれる。システム全体の流れを図5に示す。

6. 検出のアルゴリズムその一 (明らかに偽の場合)

6.1 包含関係を用いるもの

例として, 前に述べた

「人間は狼である」

という隠喩を検出する手順を説明する(図6参照)。

まず入力として以下のような state of affair を仮定

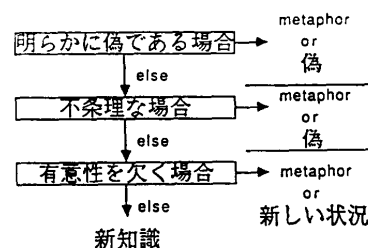


図5 システムの流れ

Fig. 5 Flow chart of system.

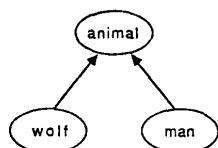


図 6 「人間は狼である」の隠喩検出
Fig. 6 Metaphor detection of "man is wolf".

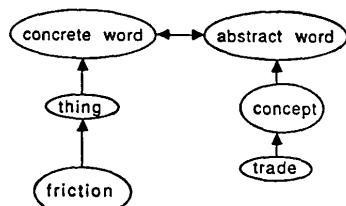


図 7 「貿易摩擦」の隠喩検出
Fig. 7 Metaphor detection of "trade friction".

する。

in s : at 1 : is, man, wolf ; yes

これを意味ネットワーク部のネットワーク部で man というノードが wolf というノードの下 (ノード間に別のノードがいくつかあってもよい) にあるかどうかで隠喩あるいは虚偽の発話のどちらかであることが検出される。この発話が虚偽の発話かどうかは、隠喩解釈を待たなければならない。この例および他のいくつかの例は計算機上で実装済みである。

6.2 属性関係を用いるもの

例として

「貿易摩擦」

の例を挙げる。これは次のように構文解析された入力形をとって入ってくる。

in s : at 1 : rub, trade ; yes

これを trade というノードの属性に rub があるかどうかを見に行き、そういう属性がなければ trade の上位のノードの属性を見に行く、こうして一番上のノードまで見に行き、そのような属性があれば普通の発話、そうでなければ隠喩あるいは虚偽の発話として検出される。図 7 のリンクの意味は 4.3 節で述べたとおりである。

7. 検出のアルゴリズムその二 (不条理な場合)

ここでも前に述べた

「彼は少し酔っている」

という例を挙げて検出の手順を説明する (図 8 参照)。

現実の発話は次のような一連の文の後に述べられた

ものとする。

juwa は酒を飲み過ぎた。

juwa は今暴れている。

それを見ていた shiro は、

「juwa はすこし酔っている」

と言った。

これらの一連の文から次のような (1)-(4) の日本語文になることを仮定する。ここで仮定された必要とされる処理としては文の時制と場所の副詞から時空位置を設定し、発話の主語や指示対象を設定することが必要とされる。

「juwa はたくさん酒を飲んだ」 (1)

「そして juwa は今暴れている」 (2)

「shiro は (1) と (2) を認識した」 (3)

「shiro は "juwa は少し酔っている" と言った」 (4)

これらの文は実際の入力として次のような状況意味論の實在的状况になることが要求される。

in s 1 : at 11 : drinking, juwa, much ; yes

in s 2 : at 12 : going-wild, juwa ; yes

in s 3 : at 11 : recognize, shiro, s 1 ; yes

in s 4 : at 12 : recognize, shiro, s 2 ; yes

in s 5 : at 12 : saying, shiro, Z ; yes

in s 6 : at 12 : gets-drunk, juwa, a-little ; yes

(11 < 12)

のような形をとり、作業領域に逐次保存される (" $<$ " は 11 よりも 12 の方が時間的に後であることを示す、Z は発話の内容に対応する)。

この一連の入力は命題の真偽チェックがなされるアルゴリズムを通過し、このアルゴリズムに入ってくる。構造部の側は

in s 1 : at 11 : drinking, A, much ; yes

in s 2 : at 12 : going-wild, A ; yes

in s 3 : at 11 : recognize, B, s 1 ; yes

in s 4 : at 12 : recognize, B, s 2 ; yes

in s 5 : at 12 : saying, B, Z ; yes

in s 6 : at 12 : gets-drunk, A, much ; yes

(11 < 12)

という抽象的状况が 4.5 節で述べた構造部の三つの部分から生成され、入力の実在的状况が入ってくるたびに抽象的状况の指数が一つ進む。構造部にはいくつかの状況が入っているがそのうちに当てはまるものに合わせて処理が進む。最後の s 6 という状況に入ってくると a-little と much が一致していないことが分

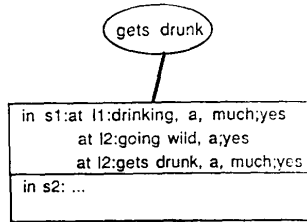


図 8 「彼は少し酔っている」の隠喩検出 (その一)
Fig. 8 Metaphor detection of "he gets drunk a little" (step 1).



図 9 「彼は少し酔っている」の隠喩検出 (その二)
Fig. 9 Metaphor detection of "he gets drunk a little" (step 2).

かる (図 9 参照)。そこで改めて a-little と much というノードを見てみるとこの二つのノードの間には対立リンクが張ってあるので隠喩あるいは虚偽の発話であることが検出される。

8. 検出のアルゴリズムその三 (有意性を欠く場合)

8.1 常識のレベルを定めるもの

ここでは常識のレベルを定めるものとして専門用語だけを取り上げる。専門用語のノードには専門用語であることが記述している。

ここでは普通の隠喩の例として、「人間は動物である」という例を挙げる。入力形は、

in s: at 1: is, man, animal; yes

という形をとる。人間の理解する順序に合わせて、アルゴリズムその一、二を通過する。最後にこのアルゴリズムに入ってきて man, animal が専門用語でないことを確認して隠喩的な発話と判定される。

8.2 諺に関するもの

焦点を各単語に定め、各単語によって呼び起こされる文脈、状況に適した諺を検索する。

例として、「隣の花は赤い」という発話を挙げる。実際の入力は、

in s: at 1: having, neighbor, flower; yes

: is, flower, red; yes

が入ってくる。アルゴリズム一、二は通過する。「隣」「花」「赤い」という語は専門用語ではないので常識のレベルにかかわらず、このアルゴリズムに入ってくる。ここで「隣」「花」「赤い」によって呼び起こされる諺を検出し、隠喩理解ルーチンに引き渡す。

9. 評 価

本論文は隠喩の検出が機械的にできるかどうかを問題としたものである。そこで現段階の評価としては定性的評価のみで十分であり、定量的な評価の段階にはない。実行速度の問題は扱わないが、小規模のシステムであるため数秒で実行は終わる。

計算機上で実際の発話に近い形から隠喩の検出を行うおうとする試みは今まで全く成されていない。これらの隠喩検出アルゴリズムでは隠喩である可能性のあるものをすべて検出する。そのため隠喩でないものも検出されてしまうが、これは文脈理解、談話理解を伴う、隠喩理解アルゴリズムにまかされることになる。

隠喩的な発話を

- (1) 包含関係に反するもの
- (2) 属性関係に反するもの
- (3) 現実の状況に反するもの
- (4) 常識のレベルに反するもの
- (5) 諺に関するもの

と分離したときそのすべてを検出する方法をこの論文は示した。

この隠喩検出アルゴリズムで分離できないものは、

- (1) 明らかに偽である場合で、虚偽の発話か隠喩的な発話かの判定をどうするか
 - (2) 不条理な場合で虚偽の発話か隠喩的な発話かの判定をどうするか
 - (3) 有意性を欠く場合で新しい状況設定か隠喩的な発話かの判定をどうするか
- のような種類の判定である。

実際に行った例の数は各々 20 ほどであるが、その本質的な特性は得られたと考えられる。

また談話理解の際に地の文章から時空位置としては現在、過去、未来、話し手の関係としては話し手の内部状態が分かればよいことが明らかとなった。

10. おわりに

本論文では検出の原理としては従来のグライスの「会話の作法」を改良したスペルベルの象徴解釈のモデルを用いた。そのデータ構造としては意味ネットワークと状況意味論を用い、実装には Prolog を用いた。これらの隠喩検出アルゴリズムでは隠喩である可能性のあるもの

- (1) 包含関係に反するもの
- (2) 属性関係に反するもの

- (3) 現実の状況に反するもの
- (4) 常識のレベルに反するもの
- (5) 諺に関するもの

をすべて検出する。隠喩でないものも検出されてしまうが、これは隠喩理解アルゴリズムにまかされることになる。隠喩検出で問題となるのは、文脈、状況と関連するところをどう扱うか、意味ネットワークの作り方に検出の性能が依存してしまうことである。意味ネットワークをうまく作ると検出できない例はまず考えられない。実際の例題でも検出が不可能な例はなかった。しかし機械翻訳などに応用することを考えると、直訳できない部分である隠喩を検出することは重要である。さらにこの隠喩の検出をいかに活かすかも隠喩理解アルゴリズムにまかされる。隠喩理解アルゴリズムは現在研究中であるが、本論文で検出されたものに対して各々の単語がどのように使われているかを具体的に調べ、単語の意味の変化をその後の意味理解に活かしていく予定である。

明らかに偽である場合で虚偽の発話かどうかの判定をどうするか、不条理な場合で抽象的状況の記述をより抽象化する方法をどうするか、諺の検出でもう少し絞り込む方法はないか、などが今後の課題として残されている。

参 考 文 献

- 1) 安井 稔：言外の意味，研究社（1978）。
- 2) 竹内晴彦：メタファー理解の試み，第33回情報処理学会全国大会論文集，3L-1（1986）。
- 3) 菅野盾樹：メタファーの記号論，勁草書房（1985）。
- 4) スベルベル，菅野盾樹（訳）：象徴表現とは何か，勁草書房（1979）。
- 5) Barwise, J. and Perry, J: *Situation and Attitudes*, MIT Press (1983).
- 6) 宮崎正弘：係り受け解析を用いた複合語の自動分割法，情報処理学会論文誌，Vol. 25, No. 6,

pp. 970-979 (1984).

- 7) 半沢幹一：たとえるものととえられるものとの関係について 記紀・風土記比喩表現の場合，日本文芸研究会研究発表（1979）。
- 8) 中村 明：比喩表現辞典，角川書店（1977）。
- 9) Brachman, R. J., Fikes, R. E. and Levesque, H. J.: *Krypton: A Functional Approach to Knowledge Representation*, *Computer*, Vol. 16, No. 10, pp. 67-74 (1983).

(昭和63年8月30日受付)

(平成元年7月18日採録)



土井 晃一（正会員）

1961年生。1984年東京大学電気工学科卒業，1988年同情報工学専攻修士課程修了，現在博士課程在学中。哲学，言語学，心理学，文化人類学を応用して，計算機による自然言語理解の研究を行っている。



田中 英彦（正会員）

昭和18年生。昭和40年東京大学工学部電子工学科卒業，昭和45年同大学院博士課程修了。工学博士。同年東京大学工学部講師，昭和46年助教授，昭和62年教授。昭和53年～54年ニューヨーク市立大学客員教授，現在に至る。計算機アーキテクチャ，並列推論マシン，知識ベース，オブジェクト指向プログラミング，分散処理，CAD，自然言語処理，等の研究を行っている。‘計算機アーキテクチャ’，‘VLSI コンピュータ I, II’，‘ソフトウェア指向アーキテクチャ’（いずれも共著），‘情報通信システム’著。電子情報通信学会，人工知能学会，日本ソフトウェア科学会，IEEE，ACM 各会員。