

プリミティブマッチング法：隠喩的見立て知識の学習法

Primitive Matching Approach : A Method for Eliciting Metaphorical Knowledge about Viewing

諏訪 正樹* 元田 浩*
Masaki Suwa Hiroshi Motoda

* (株)日立製作所 基礎研究所
Advanced Research Laboratory, Hitachi Ltd., Hatoyama, Saitama 350-03, Japan.

1993年4月9日 受理

Keywords: analogical mapping, metaphor interpretation, primitives, knowledge about view-relations.

Summary

Cross-modal metaphors are based on some “view-relations” between concepts in different domains, i.e. *knowledge about viewing a mental (or abstract) concept as another physical concept*. Our goal is to establish an effective method of computationally eliciting those pieces of knowledge from cross-modal metaphors we human beings use in everyday life. Since we understand that analogical mappings between two non-identical concepts underlying cross-modal metaphors represent viewing-relations, we proposed a method for determining analogical mappings of metaphors as follows. Basically, we employ the *structural consistency* constraint in determining analogical mappings which allows the mappings of *non-identical* two relational predicates. In general, however, only this simple version of structural consistency is too weak as a constraint and tends to produce so many meaningless alternatives of analogical mappings. This paper addresses the problem by employing a *semantic* constraint, called *primitive matching*, only at the initial stage of finding analogical mapping; based on the representation that verbal concepts are described by use of abstract primitives only, the program first searches for the pairs of *first-order* literals with *exactly the same primitive predicates* which yield some initial mappings. And after that, the program searches for the whole correspondence between two concepts which does not contradict the initial mappings, based on the simple version of structure consistency constraint. In this paper, we made an experiment of analyzing some cross-modal verbal metaphors by the primitive matching approach, as a result of which we show that the approach proposed is more advantageous in cpu-time saving than the simple version of structural consistency approach in some classes of verbal metaphors, and we also illustrate the obtained knowledge about viewing mental concepts as physical ones, which is general enough for use in interpreting other metaphoric expressions.

1. はじめに

ある概念(ターゲット概念)を別の概念(ソース概念)で見立てることによって、前者のある特徴をきわだたせるという隠喩的手法は、人間が物事を考えまたは表現するうえで中心的役割を果たしている[Carbonell 80, Lakoff 80]. Barnden は、我々人間は心的概念を物理的概念に例えること(つまり、隠喩的表現を使うこと)なしには心の状態を十分に表現できないと主張

している[Barnden 89]. 彼の主張によれば、隠喩文は心的概念を物理的概念に見立てる見立て知識、すなわち人間固有の「心的概念の物理的イメージ」の宝庫であるといえる。本稿の目的は、隠喩文を計算機で処理することによって、隠喩文から「心的概念の物理的イメージ」を抽出することを通じて、我々人間が行う類似性認識の源を探ることにある。

隠喩理解のための計算モデルとして、我々は、以下のアプローチをとる。隠喩文を支える異なる二つの概念(ソース概念とターゲット概念)の意味記述間の写

像を求め、その写像をもって隠喩を構成する類比 (analogical mapping) とみなす*1, という手法である。これは一般に類推的アプローチと称される手法である [Martin 90], [Carbonell 80, Gentner 83] などがその代表例である。我々が類推的アプローチを採用する理由は、隠喩文に潜む「異なる概念間に成立する類比関係」こそ心的概念と物理的概念の隠喩的結びつきに関する知識であると考えからである。

類推的アプローチに属する研究の大半は、二つの意味記述間の類比を決定する際にかなる制約 (または知識) を使うかに関するものである。Holyoak の分析 [Holyoak 89] によると、それらの制約は次の3種類に大別できる。構造的な同型性制約 (structural consistency), 意味的類似性に関する知識 (semantic similarity), 運用的制約 (pragmatic centrality) である。諏訪らは、隠喩文処理におけるこれらの制約の適用性に関して検討し、以下の洞察を得た [Suwa 91, 諏訪 93]。概念木 (abstraction hierarchy) などの意味的類似性知識だけで隠喩文に埋め込まれている隠喩的結びつきを発見することは不可能であり、基本的には構造的な同型性制約を用いるのがよいが、そのみでは類比の発見に多大の計算時間がかかって好ましくない (2章で詳説)。また、自然言語文の意味決定が話者の意図や (話す) 目的などの運用的制約 (文脈) に左右されることは間違いのない事実で、隠喩文もその例外ではないが、隠喩文の場合、発話者の意図とは、ソース概念に言及することによってターゲット概念からは通常連想しないような性質を顕現化することである。しかし、それは隠喩文理解の結果として発見すべきことであって、この種の運用的制約を隠喩文処理のための知識として与えておくことは好ましくない。Gentner も同様の主張を行っている [Gentner 88]。比喩文が一文だけ単独で与えられた場合 (この場合、文脈情報はない) にでも我々人間は比喩文を理解できることを考慮し、運用的情報なしで比喩を理解する手法を提供しなければならないという主張である。

これらの検討の結果、諏訪らは、構造的な同型性制約を基本にして、プリミティブマッチングという意味的制約を併用する手法 (プリミティブマッチング法—PM法と称する) を提案した [Suwa 91]。Gentner の

structure-mapping 理論 [Gentner 83] との相違点、および意味的制約を用いることの必然性に関しては、2章で詳しく述べる。PM法は、ソース概念とターゲット概念の間に構造的な類比関係を持つような隠喩文、つまり、動詞隠喩文 (verbal metaphor) を代表例とする構造比喩 (もしくは関係比喩) を処理するための手法であり、名詞・形容詞句 (metaphor of adjective noun combination) を代表とする属性比喩*2 は本稿では扱わない。

本稿では、PM法を隠喩理解の計算モデルとしてインプリメントし、いくつかの動詞隠喩文を例にして、PM法および構造的な同型性制約のみを使った手法 (SM法と呼ぶ) の性能を比較するための基礎実験を行った。3章で、計算モデルのアルゴリズムを詳説する。4・1節でPM法の類比決定性能に関する実験結果について議論し、4・2節で動詞隠喩文から抽出した「心的概念の物理的イメージ」に関する知識を例証する。

2. プリミティブマッチング法

2・1 基本コンセプト

PM法は、構造的な同型性制約および意味的制約の両者を併用して、ソース概念・ターゲット概念間の写像を決定する。

我々は、構造的な同型性制約の概念として、[Holyoak 89] の定義を採用する*3。関係述語が1個以上のオブジェクトを引数として持つリテラルを1階命題と呼び、高階関係述語が1個以上のオブジェクトまたは1階命題を引数として持つリテラルを高階命題と呼ぶ。ソースおよびターゲット領域間に写像 map が存在し、

$$map(o_{t,i}) = o_{s,i} \quad \text{および} \quad map(R_{t,i}) = R_{s,i}$$

を満たすとする。ただし、 $o_{t,i}$, $o_{s,i}$ は、ターゲット領域、ソース領域での i 番目のオブジェクトを表し、 $R_{t,i}$, $R_{s,i}$ はターゲット領域、ソース領域での i 番目の関係述語を表す。このとき、ターゲット側で $R_{t,i}(o_{t,1}, \dots, o_{t,n})$ という命題が成り立ち、ソース側でも $R_{s,i}(o_{s,1}, \dots, o_{s,n})$ が成り立つならば、この写像 map は構造的な同型性制約を満たす写像であるという。この性質を満足する写像は非常に広い概念であり、同型でありさえすれば、写像により類比づけられるオブジェクトどうし、関係述語どうしが異なるものであっても構わない。Gentner の structure-mapping 理論 [Gentner 83] は、Holyoak の定義の構造的な同型性写像に次の3つの制約を付加したものである。1引数の1階命題の写像を許さないこと、互いに引数の数が同じであるリテラルに関して、もし関係述語が異なるならばそれら

*1 以下では、写像と類比を全く同じ意味で使用する。

*2 構造比喩 (または関係比喩) vs 属性比喩という分類は、[Gentner 89] より引用した ([諏訪 93] を参照のこと)。

*3 Holyoak はこの定義の制約を提唱したのではなく、過去の文献で提唱されているさまざまなバージョンの構造的な同型性制約の基本的共通部分を一般化し、総括としてこの定義に言及しているにすぎない。

の写像を許さないこと、1階命題よりは高階命題の写像を優先すること**の3点である。

Holyoak の定義は、単純で拘束条件としては弱いために、同型性を満たすような写像は一般に数多く生成可能であり、すべての場合を計算するコストは莫大なものになる。Gentner の structure-mapping 理論はこの点を解決する強い制約になっているが、互いに異なる関係述語の写像を許さないという性質は、言語の類似性判定に不向きであり、また、異なる関係述語の類比を「心的概念の物理的イメージ」とみなし、隠喩文から抽出するという我々の目的に全くそぐわない。そこで、我々は、Holyoak の定義の構造的な同型性制約を基本にして、さらに、以下に示す意味的制約を拘束条件として付加することにした。

まず、ソース、ターゲット両領域の意味記述のなかから、1階の関係述語が同じ (identical) である命題を探し、その関係述語どうし、引数のオブジェクトどうしを類比づける写像を生成する。そして、次に、その写像と1対1対応原則に照らして矛盾しないような両領域間の写像を、構造的な同型性制約により発見する。この後半プロセスでは、異なる関係述語の写像も許す。この点が Gentner の理論との本質的な違いである。両プロセスを経て生成される写像は一般に複数通り存在するが、そのうち最も類比の数の多いものを正解の類比として選ぶ。そして、後半プロセスで生成される異なる述語間の類比を、隠喩文を支える見立て知識として抽出する。前半プロセスで生成された写像が、初期拘束条件として両領域間で類比づけ可能な高階関係を求める際にむだな探索を未然に枝刈りし、計算コストを軽減するのに有効であると期待できる。

本手法がうまく機能するか否かは、前半プロセスにおいてどれだけ拘束力の強い写像が生成できるかにかかっている。一般に、異なる領域の概念は異なる関係述語で意味記述されることが多いため、両領域中に関係述語が一致する命題が存在すること自体が非常にまれであり、そのため前半プロセスが機能しないことが考えられる。そこで、我々は、概念の意味を記述するための基本言語として、あらかじめ有限個の意味素集合を考え、前半プロセスを、「1階の意味素述語が一致する1階命題を探すプロセス」と設定した。以下、前半プロセスをプリミティブマッチングプロセスと呼ぶ。

プリミティブマッチングプロセスを1階の述語の一致のみに限ったのは、以下の三つの理由による。

- ・高階関係の意味素とは因果関係や動機などを記述するための述語であり (2・2 節参照のこと)、どの

領域の動詞概念を記述する際にも意味記述中にしばしば現れるため、高階関係の意味素の一致は後半プロセスへの強い拘束条件にはなり得ないこと。

- ・なかでも、因果関係 (例えば cause) は意味記述中に複数現れる場合が多く、二つの動詞概念間の写像を決定する際、ある一つの因果関係述語 (cause) を相手概念のどの因果関係述語 (cause) と類比づけるかで場合分けが生じ、あまり計算コストの低減が見込めないこと。
- ・高階関係述語の意味素の一致は、大局的な構造の類似性を示すものであり、意味的な類似性を反映する制約にはならないため、

である。

2・2 意味素の収集

我々は、約 200 の動詞概念の意味記述を行った経験から、図 1 に示す意味素を収集した。意味素は、(1)オブジェクトの属性、(2)オブジェクトの状態を表す述語、(3)オブジェクトに関する動作・作用を表す述語、(4)オブジェクトどうしの関係を表す述語、(5)高階関係 (因果関係を含む) を表す述語、に分類できる。(1)がファンクタを表す意味素、(2)~(4)が1階の意味素、(5)が高階の意味素である。

動詞概念を意味表現するためには、あるオブジェクトの動作や作用や関係を表す意味素以外に、状態を表す意味素が本質的に必要になる場合がある。状態を表す意味素のなかには、品詞分類上の形容詞的な概念が含まれる。しかし、形容詞的概念を意味表現するためにどのような意味素を用意すればよいかは、それ自体重要で非常に難しい課題であるため[内海 93a]、ここ

属性等のファンクタ

the function of, the quantity of, the nature of, the sum of, the difference of, the motion of, the direction of, the location of, the distance of

1階関係の述語

状態

exist, be located, transfer, appear, disappear, change, rotate, traverse, stay, good, bad, happy, unhappy, large, small, strong, weak, true, false

関係

include, intersect, touch, located apart, same, different, ...er, ...est, a part of

動作、作用

possess, do, proceed, operate power, hold, use, live, mean, attention, like, want, know, think

高階関係の述語

when, where, cause(result), motivate, the purpose of, the reason why, how, make ... do, equivalent to ... (relations) disappear, appear

図 1 意味素集合

* 4 これを Gentner は, systematicity principle と呼んでいる。

では詳しくは触れない。そこで、我々は、楠見の研究 [楠見 89] から快不快, 善悪, 大小, 強弱などが形容詞的概念を分類する際の基本軸になり得る点に着目し(5章参照), それらに対応する基本的形容詞のみを意味素として与えた。

意味素の収集は以下に示す方法で行った。まず、動詞の意味を英英辞典で調べる。我々は Collins COBUILD 辞典 [COBUILD 87] を用いた。一般に言葉の意味は多義的であるが、そのなかで最も語源的に中心的な意味を選ぶ。選んだ意味表現のなかに使われている言葉を、また英英辞典で調べる。このように再帰的に辞書を引いていくうちに、ある言葉を表現するために似た意味の別の言葉を使用している状況が現れれば、その二つの言葉の上位概念としてふさわしい言葉を考えだし、それを意味素とする。

2.3 意味素による意味記述法

我々は、意味素だけを用いて約 200 個の動詞概念の意味記述を行った。意味として何を記述し何を記述しないかを一般的に決めることは、文脈依存性という言語の特質上不可能であり、記述する人の感性・主観・体験に左右されることは避けられない。内海らは、形容詞の記述において、意味を、場面や状況に影響されない客観的な部分としての「意味の核」と、主観的な色合いの濃い「意味の肉」に分けて考えている。本手法では、動詞概念の「意味の核」に類する記述のみを行うことにする。

記述における規範は、以下の三つである。

1. オブジェクトの属性や関係を表す言葉として、意味素だけを使用すること。
2. 動詞概念が表す行動や関係そのものの記述にとどまらず、その行動や関係が生じる契機や前提、あるいはその行動や関係の波及効果や結果を、高階関係の意味素で記述すること。しかし、例えば因果関係や波及効果としてどの範囲まで書くべきかを示す明確な指針はない。あくまでも、「意味の核」に相当する部分の記述にとどめる。
3. 登場するオブジェクトが、名詞概念木 (taxonomy) 中のどの分類に属するものかを記述する。実例を以下に示す。本稿では、

“He shot down my opinion in the argument”.
 という動詞隠喩文を議論に使用する。shoot down が隠

* 5 1階の意味素述語 p および p' に関して、 $p=p'$ が成り立つリテラルを探すプロセスなので、同じ引数位置にあるオブジェクトどうしを類比づけることとなる。したがって、 $\langle t_i, t_i \rangle$ を新たに類比に加える。高階構造の類比プロセス、異なる 1階関係どうしの類比生成プロセスと比較のこと。

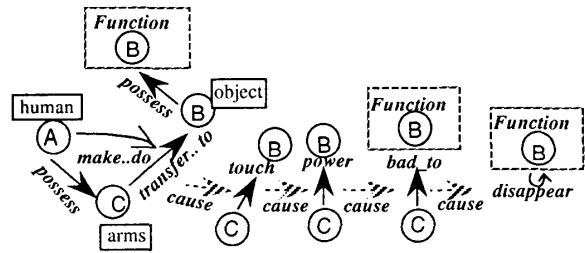


図 2 意味素を用いた shoot down の意味記述

喩的に使われており、最も近い意味の動詞は criticize である。我々は、“A shoots down B” の意味構造を意味素を用いて図 2 のように記述した。

「A が所有 (possess) しているもの C を、B に向けて移動 (transfer-to) させることに加担 (make...do) し、それが原因 (cause) で C が B に接触 (touch) し、それが原因 (cause) で C から B に対して力が働き (power), それが原因 (cause) で B の機能 (function-of) にとって悪い (bad-to) 影響を及ぼし、それが原因 (cause) で B の機能 (function-of) が消滅 (disappear) する。」

と読む。下線で示した言葉が意味素である。

3. PM 法アルゴリズム

PM 法による計算モデルは、二つの動詞の意味記述を入力とし、両記述間に成り立つすべての類比を生成する。

D および D' を二つの動詞の意味記述式集合とする。 hr, h, a, p, t を、それぞれ、高階関係式、高階関係の述語記号、1階関係式、1階関係の述語記号、オブジェクトを表す項とする。このとき、類比 ϕ は、項、1階関係式、高階関係式と 1対1 の対応関係の集合で定義される。

PM 法では以下の三つのステップを経て、意味記述間の類比を決定する。

1. (1階意味素のマッチング)

以下のルールに従って、1階の意味素述語の一致するリテラルを探す。1階関係式 $a=p(t_1, \dots, t_n) \in D$ および $a'=p'(t'_1, \dots, t'_n) \in D'$ と現在の類比 ϕ に関して、もし $p=p'$ で、 $\{\langle a, a' \rangle\} \cup \{\langle t_i, t'_i \rangle \mid i=1, \dots, n\} \cup \phi$ が 1対1 対応を満たすなら、類比 ϕ を以下のように更新する。 $\phi = \phi \cup \{\langle a, a' \rangle\} \cup \{\langle t_i, t'_i \rangle \mid i=1, \dots, n\}$ *5。このプロセスを繰り返して最終的に得られる類比を ϕ_1 とする。この ϕ_1 が、以下の 2 プロセスで類比を求める際の拘束条件として働く。

2. (高階関係式の類比)

ϕ_1 と矛盾しないような高階関係どうしの類比を探す。 $\alpha_i = p_i(t_{i,1}, \dots, t_{i,m(i)})$, $\alpha'_i = p'_i(t'_{i,1}, \dots, t'_{i,m(i)})$ とする。高階関係式 $hr = h(\alpha_1, \dots, \alpha_n) \in D$ および $hr' = h'(\alpha'_1, \dots, \alpha'_n) \in D'$ と、現在の類比 ϕ に関して、もし、すべての i ($1 \leq i \leq n$) に対して、1対1対応類比 $\phi_i = \{\langle \alpha_i, \alpha'_i \rangle\} \cup \{\langle t_{i,j}, t'_{i,k(j)} \rangle \mid j=1, \dots, m(i)\}$ が存在し*6、しかも、 $(\cup_i \phi_i) \cup \phi$ が1対1対応という条件を満たすなら、類比 ϕ を以下のように更新する。 $\phi = \phi \cup \{\langle hr, hr' \rangle\} \cup (\cup_i \phi_i)$ 。このプロセスを繰り返して最終的に得られる類比を ϕ_2 とする。

3. (異なる1階関係どうしの類比)

述語記号は異なっても、引数がすでに ϕ_2 によって類比づけられている1階関係どうしを、以下のルールに従って類比づける。 $\langle a, a' \rangle \notin \phi_2$ を満たすような1階関係式 $\alpha = p(t_1, \dots, t_n) \in D$ および $\alpha' = p'(t'_1, \dots, t'_n) \in D'$ に関して、もし、 $\{\langle t_i, t'_{i(i)} \rangle \mid i=1, \dots, n\} \subset \phi_2$ なら、類比 ϕ を以下のように更新する。 $\phi = \phi \cup \{\langle \alpha, \alpha' \rangle\}$ 。このプロセスを繰り返して最終的に得られる類比を ϕ_{total} とする。

一般に、プリミティブマッチングプロセスでは複数通りの類比 ϕ_1 が得られ、そのおのおのに関して ϕ_{total} が生成される。そのなかで、最も類比対応関係の数が多きものを二つの意味記述間の正解の類比として選択する。

例文 “He shot down my opinion in the argument”.

において、上記のアルゴリズムに従いソース概念 ‘shoot down’ に対応する適切なターゲット概念を発見するという問題を考えよう。主語に ‘human being’ をとり、目的語に ‘opinion’ をとるような動詞がターゲット概念の候補である。例えば、‘refuse’, ‘agree’, ‘disagree’, ‘listen to’, ‘neglect’, ‘understand’, ‘insist’ などの動詞があげられる。それぞれの候補に関して ‘shoot down’ との間に類比を求め、対応関係の数が最大の類比 (最大類比) を生成した候補動詞をターゲット概念として選ぶ。上記の例文の場合、‘criticize’ が最大類比を生成する。ここでは、‘shoot down’ と ‘criticize’ の間に生成される類比を例としてあげる。 “A’

* 6 互いに異なる関係述語 p, p' を持つような1階命題 α_i, α'_i を類比づける場合には、異なる位置にある引数どうしを類比づけることを許すことにする。ただし、本文に示すように、類比づけようとする1階命題の引数の個数 ($m(i)$ 個) が一致していること、およびすべての引数に関して (すべての j に対して) 1対1対応の類比が存在することが条件である。異なる1階関係どうしの類比生成プロセスでも、同様の類比を許す。

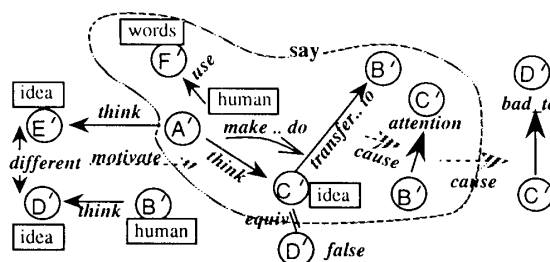


図3 意味素を用いた criticize の意味記述

$\phi_{total} = \{\langle c,c' \rangle, \langle b,b' \rangle, \langle a,a' \rangle, \langle \text{function_of}(b), d' \rangle, \langle \text{transfer_to}(c,b), \text{transfer_to}(c',b') \rangle, \langle \text{bad_to}(c, \text{function_of}(b)), \text{bad_to}(c',d') \rangle, \langle \text{make_do}(a, \text{transfer_to}(c,b)), \text{make_do}(a', \text{transfer_to}(c',b')) \rangle, \langle \text{touch}(c,b), \text{attention}(b',c') \rangle, \langle \text{possess}(a,c), \text{think}(a',c') \rangle, \langle \text{possess}(b, \text{function_of}(b)), \text{think}(b',d') \rangle\}$

$\phi_{total} = \{\langle c,c' \rangle, \langle b,b' \rangle, \langle a,a' \rangle, \langle \text{function_of}(b), d' \rangle, \langle \text{transfer_to}(c,b), \text{transfer_to}(c',b') \rangle, \langle \text{bad_to}(c, \text{function_of}(b)), \text{bad_to}(c',d') \rangle, \langle \text{make_do}(a, \text{transfer_to}(c,b)), \text{make_do}(a', \text{transfer_to}(c',b')) \rangle, \langle \text{power_op}(c,b), \text{attention}(b',c') \rangle, \langle \text{possess}(a,c), \text{think}(a',c') \rangle, \langle \text{possess}(b, \text{function_of}(b)), \text{think}(b',d') \rangle\}$

In both cases, $\phi_1 = \{\langle c,c' \rangle, \langle b,b' \rangle, \langle \text{function}(b), d' \rangle, \langle \text{transfer_to}(c,b), \text{transfer_to}(c',b') \rangle, \langle \text{bad_to}(c, \text{function}(b)), \text{bad_to}(c',d') \rangle\}$

但し、高階関係述語に関する類比は省略する

図4 shoot down と criticize の意味記述間の類比

criticizes D'” の意味構造を図3に示す。PM法で ‘shoot down’ と ‘criticize’ の類比を求めると、図4の二つの類比が得られる。プリミティブマッチングプロセスで得られる類比 ϕ_1 はただ1種類であるが、その後のプロセスで2通りに場合分けされる。両者は、attention (b, c) が類比づけられる相手が touch (c, b) か power_op (c, b) かの違いである。両者とも対応づけの数が同じなので、ここでは、両方とも正解の類比であるとする。

4. 実験結果

4.1 PM法による計算コスト

我々は、Lakoffの著書[Lakoff 80]および他の文献から、動詞が隠喩的に使われている17個の隠喩文を集め、Holyoakの定義の構造的な同型性制約だけによる類比決定法(SM法と呼ぶ)に比べて、PM法が計算コストの面でどれだけ有利かを調べた。おのおのの隠喩文に関して、文中で隠喩的に使われている動詞(ソース概念)に対応するターゲット概念の動詞は計算モデルに与えて、両者の意味記述間のすべての類比を算出するのに要するCPU時間を測った。

図5に、実験に使った17個の隠喩文を示す。おのおのに付記した括弧内の動詞が、それぞれに与えたターゲット概念動詞である。表1に、PM法およびSM法で要した計算コストと、PM法によるコストゲイン値(PM法により何倍速くなったか)をまとめた。17のケ

Case 1:	We need to <u>explore</u> the issues carefully.	(examine)
Case 2:	He tried to <u>grind out</u> new solutions.	(elaborate)
Case 3:	If you use that strategy, he'll <u>wipe</u> you out.	(destroy)
Case 4:	The great idea was <u>planted</u> in his youth.	(developed)
Case 5:	He <u>shot down</u> my opinion in the argument.	(criticize)
Case 6:	Where'd you <u>dig up</u> that idea ?	(discover)
Case 7:	They <u>exploded</u> his latest theory. (disprove)	
Case 8:	I just can't <u>swallow</u> that claim. (accept)	
Case 9:	He <u>pointed out</u> the demerits of the strategy.	(show)
Case 10:	Tigers <u>crushed</u> Giants in the baseball game.	(defeat)
Case 11:	The argument <u>collapsed</u> .	(fail)
Case 12:	She <u>killed</u> a process. (stop)	
Case 13:	I <u>got out of</u> a lisp. (stop doing)	
Case 14:	We need to let that idea <u>percolate</u> . (spread)	
Case 15:	There are too many facts here for me to <u>digest</u> them all. (understand)	
Case 16:	His words <u>carry</u> little meaning. (represent)	
Case 17:	Try to <u>pack</u> more thought <u>into</u> fewer words. (express by)	

図5 実験に使用した動詞隠喩文

表1 類比生成の計算コスト (PM法 vs SM法)

Group	Case No.	PM (msec)	SM (msec)	Cost Gain (SM/PM)
1	1	949	5684	6.0
	2	7130	39350	5.5
	3	1351	5134	3.8
	4	2351	8585	3.7
	5	2615	8451	3.2
	6	299	868	2.9
	7	4984	15452	3.1
	8	734	1350	1.8
	9	417	500	1.2
	10	217	284	1.3
	11	3168	2300	0.73
2	12	67	68	1.0
	13	83	100	1.2
	14	550	450	0.82
	15	367	350	0.95
	16	567	450	0.79
	17	832	837	1.0

ースは、両動詞の意味記述に共通の1階の意味素が存在した隠喩文 (Group 1) と、存在しなかった隠喩文 (Group 2) に分類できる。Group 2 では、PM法はSM法に比べてわずかではあるが多くの計算コストを要する。それは、共通の1階の意味素が存在しないので、そのマッチングを探すプロセスがむだに終わり、その後のプロセスではSM法と全く同じプロセスをたどることになるからである。一方、Group 1 では、ほとんどの隠喩文でPM法による効果が見られる。その理

*7 厳密に言えば、探索空間の大きさは全部の項の数どうしを掛け合わせた量とするのが正しいが、本アルゴリズムでは、シンタクティックにマッチング不可能なリテラルどうし (例えば、1引数の1階関係と、2引数の1階関係) は、そもそもマッチング可能かどうかの探索を行わないようにしているので、実効的には、上記の式のように、シンタクティックにマッチング可能なペアどうしの積をとり、それらの総和を求めるといった計算で十分である。

由は、高階関係式どうしの類比づけが可能かどうかを調べるときに、1階の意味素のマッチングにより得られた類比 (アルゴリズム中の ϕ_1) との矛盾チェックを行うことによって、高階関係式の内部構造にまで深く調べなくても簡単に類比づけの可否を判定できる、すなわち、全体の類比を求める際の探索空間が枝刈りされるからである。

1階の意味素のマッチングによる探索空間の減少度合いとCPU時間コストゲインの関係を明確にするために、両動詞の意味記述の類比づけ複雑性 (mapping complexity と呼ぶ) という新しい指標を導入する。mapping complexity $MC(v1, v2, \phi)$ は、動詞 $v1$ と $v2$ の類比を求めるプロセスの途中において、現在までに得られている類比が ϕ であるとしたとき、どのくらいの類比づけの複雑性が残っているかを大まかに表す指標であり、次式に従い計算する。

$$MC(v1, v2, \phi) = N_{f,1}(v1, \phi) * N_{f,1}(v2, \phi) + N_{f,2}(v1, \phi) * N_{f,2}(v2, \phi) + N_{f,n}(v1, \phi) * N_{f,n}(v2, \phi) + N_{h,1}(v1, \phi) * N_{h,1}(v2, \phi) + N_{h,2}(v1, \phi) * N_{h,2}(v2, \phi) + N_{h,n}(v1, \phi) * N_{h,n}(v2, \phi)$$

ただし、 $N_{flag,i}(v, \phi)$ ($flag=f$ or $h, i=1$ or 2 or n) は、動詞 v の意味記述中において類比 ϕ の要素として現れないリテラルの数を表す。第1下つき文字 $flag$ が f の場合は1階関係式を示し、 h の場合は高階関係式を示す。また、第2下つき文字 i は引数の個数を表す (ただし、 $i=n$ の場合に限り、引数が三つ以上の関係式を示す)。例えば、 $N_{f,1}(v, \phi)$ は、動詞 v の意味記述中において類比 ϕ の要素として現れないリテラルで1引数の1階関係式の数を示す。また、 $N_{h,n}(v, \phi)$ は、動詞 v の意味記述中において類比 ϕ の要素として現れないリテラルで引数が三つ以上の高階関係式の数を示す。したがって、 $MC(v1, v2, \phi)$ は二つの動詞の意味記述において残っている、構造的に類比づけ可能かどうかを調べなければならない1階および高階関係式の組合せの数を表す。換言すれば、 ϕ という部分類比が得られている段階において全体類比決定のために残された探索空間の大きさを意味している*7。

1階の意味素のマッチングのあとに残っている mapping complexity の割合 ($MC(v1, v2, \phi)/MC(v1, v2, \{\})$) を横軸に、PM法によるコストゲインを縦軸にとり、Group 1 に属する11のケースについて両者の関係を調べた結果が図6である。ただし、 ϕ_1 は1階の意味素のマッチングにより得られた部分類比である。二

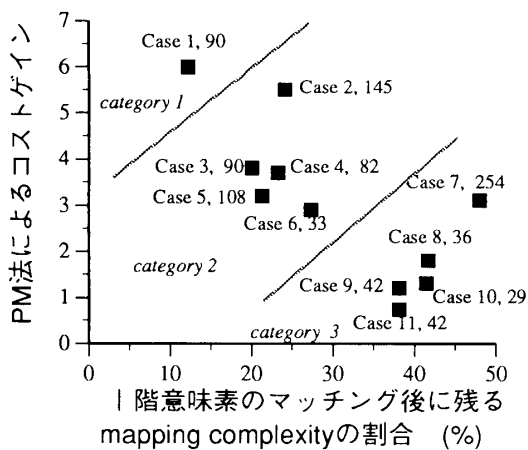


図6 PM法による計算コストの軽減度合い

つの動詞の意味記述のなかにいくつの1階の意味素のマッチングが存在するかによって、11のケースを三つのクラスに分類して、図6の結果を考察する。Category 1は、三つのマッチングが存在する場合でCase 1がそれに属する。Category 2は、二つのマッチングが存在する場合で、Case 2からCase 6がそれに属する。Category 3は、一つのマッチングが存在する場合で、Case 7からCase 11がそれに属する。意味素のマッチングの個数が多いほど、それによって枝刈りされる探索空間が大きくなり、その結果PM法がSM法に比べてより有利になるという傾向が見られる。この結果は、意味素のマッチングにより構造的同型性の発見プロセスにおけるむだな探索が避けられることを例証するものである。

次に、カテゴリー内でのコストゲインの比較について考察する。図6のおおのこのプロットには、それぞれのケースの類比決定の matching complexity 値 (MC(v1, v2, {})) が付記されている。それは、ほぼ二つの動詞の意味の記述量に比例する値である。Category 2および3において、記述量が比較的大きい場合にはPM法がSM法より優れているという傾向が見られる。

4・2 心的概念の物理的イメージ

本稿で使用した意味素は、'attention', 'think', 'know', 'mean'などの心的概念を表すものと、'power_op', 'touch', 'located'などの物理的概念を表すものに大別できる。隠喩文を成立させている類比を決定し、異なる意味素間の類比関係を「心的概念の物理的イメージ」に関する人間固有の知識として抽出した。

図7に、17個の隠喩文から抽出した 'attention', 'think', 'know' の物理的イメージに関する知識を示す。図中の矢印が、それぞれ、心的概念（長方形で囲まれたもの）の物理的イメージを表している。それぞれの心的概念の物理的イメージは、その概念述語がど

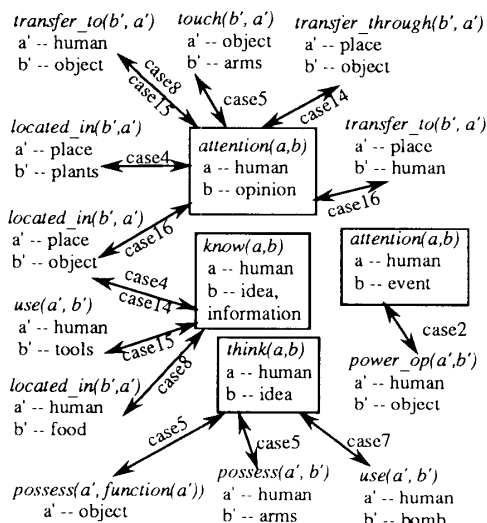


図7 心的概念の物理的イメージ (隠喩文から抽出)

ういう種類のオブジェクトを引数にとるかに依存する。例えば、'attention'の場合、第2引数に 'event' をとる場合、その 'event' は 'object' であると見立てられ、関係式 'attention (human, event)' は関係式 'power_op(human, object)' に見立てられる。つまり、「人が何かの 'event' に気づくとは、すなわち、何かの 'object' に外力を及ぼすことに等しい」という隠喩的知識である。また、'attention' が第2引数に 'opinion' か 'idea' をとる場合、'opinion' (または 'idea')は、'object' もしくは 'human' に見立てられ、'attention' の第1引数である 'human' は 'places' に見立てられる。これは、「人間がある意見や考え方に気がつくとは、すなわち、その意見や考え方があたかも物体のようにその人のいる場所へ移動してくる ('transfer_to') もしくはその場所を通過し ('transfer_through')、何らかの相互作用を起こす ('touch') かその場所にとどまる ('located_in') ことに等しい」という隠喩的知識である。

5. 関連研究および議論

Gentner, および Falkenheiner らの structure-mapping engine[Falkenheiner 89]とPM法の本質的な違いは、PM法では異なる関係述語の類比を許す点にある。異なる概念間の比喩的結びつきに関する知識を抽出する目的上、この制限緩和はぜひ必要であった。しかし、この緩和により二つの意味記述間に生成できる類比は膨大な数になる。本論文中でのSM法がこれに該当する。この点を解決するために、本研究では、類比決定の初期段階においてのみ類比を1階の意味素の一致に限定するという意味的制約を設け、それが無意味な類比の生成を抑え計算時間を低減する役目を果

たしているといえる。

隠喩文から抽出した「心的概念の物理的イメージ」に関する知識は、Norvig が文章理解プログラムに与えた view-link 知識[Norvig 86]よりも一般性が高い。view-link は、「kill は defeat-convincingly に見立てることができる」という例に見られるように、特定のより具体的な概念間の関係を示す知識であるのに対し、「心的概念の物理的イメージ」は意味素間の比喩的結びつきを示す知識なのでより一般的である。例えば、view-link がなぜ成立するのかを説明する根拠になり得る一般性を有している。kill および defeat-convincingly は、「悪い・方向へ・変化・させる」*8 という意味素レベルで互いに「意味の核」が関係づけられており、それが両者を view-link として結びつける根拠と考えられる。

内海は、属性隠喩文の理解の基盤として、形容詞の意味を有限個の意味素で表現する手法を提案している点[内海 93a, 内海 93b]で、本研究と同じ考え方に基づいているが、以下の点で両者は本質的に異なる。内海は、意味素間の連想関係の知識（ A , B を意味素とすると、 A から B が連想できるという知識を $A \rightarrow B$ と表す）を与えているのに対し、本研究では、意味素間の比喩的な連想関係の学習を目的とし、その学習メカニズムを提案している。隠喩を構成する各概念の意味記述中で、すでに意味素レベルで一致している部分から生成される類比写像を拘束条件にして、異なる意味素間の類比を求めるという研究例は、過去に存在しない。

楠見らは、多数の概念を統計的手法により分析した結果、概念の意味空間が、快不快や善悪などの「評価」軸、大小や強弱などの「力量性」軸、変化や動きなどの「活動性」軸により構成される 3 次元空間として表現できることを見出した[楠見 89]。本研究の最終目的は、実際の隠喩文を支える意味素間の比喩的關係を多数抽出することを通じて、意味素間の類似性に関する洞察を得ること、そして、その結果と楠見の距離空間との関係を探ることである。

PM 法の隠喩解釈能力について議論する。隠喩理解システムを構築する場合、そのシステムが隠喩を正しく解釈して正しい類比関係を出力できるかを評価しなければならない。しかし、何をもって正しい類比と判定するかを意味的に決定できるような尺度を用意することは現実上極めて難しいため、PM 法では、「最も対応づけの個数の多いものを正解とする」というシクタ

* 8 「悪い・方向へ・変化・させる」の・で結ばれたおのおの言葉が意味素である。

クティックな規則だけを用いている。したがって、‘shoot down’ と ‘criticize’ の例のように、最終的に対応づけの個数が最大の類比が複数存在した場合には(図 4)、そのうちどれを選ぶかは PM 法では決定できない。隠喩文を自動的に計算機に理解させることを目的にしたシステムを構築するのであれば、上記の選択を可能にするような知識を与えることが必須の課題となるが、本稿は、隠喩文に暗示的に含まれていて、我々人間が無意識のうちに利用している「意味素間の類比関係」(4・2 節で示したもの)を抽出することを目的にしているため、システムが完全に自動的に正解の隠喩解釈を出力できなくてもよいという立場をとっている。つまり、上記の単純な規則だけを用いて正解の候補を数少なく絞った後は、ユーザがそのなかから正解を選択し、システムは選択された類比から「意味素間の類比関係」に関する知識を獲得し、蓄えておくという利用形態が考えられる。正解の候補を絞るという観点からすれば、PM 法における「意味素マッチング」という意味的制約と、「最も対応づけの個数の多いものを正解とする」という規則で十分であると考えられる。

6. 将来の課題

本手法に基づいて、より多くの隠喩文を解析し意味素間の比喩的關係に関する知識を抽出することを通じて、概念の類似性に関する洞察を得ることが、将来の課題である。そのために必要なより短期的な課題を以下に示す。

まず、本手法を増進的 (incremental) な学習システムに拡張し、学習した知識を他の隠喩文を解析する際に使用できるように改善することが、必須の課題である。1 階の意味素の一致という意味的制約は非常に強力である反面、両領域中に共通の 1 階の意味素が存在しない場合には本手法が全く機能しないことは、実験の項目で述べたとおりである。そこで、意味素の一致という制約を緩和して、学習により得られた意味素間の比喩的結びつき知識をも意味的制約として使ってよいというように拡張すれば、より多くの隠喩文の解析を本手法で行うことができる。

しかし、学習すればするほど、意味的制約として使える知識が増え、結果的に生成される類比の場合の数は膨大となる。これは学習システムにおいて一般的に発生する学習パラドックスであり、それに対処するためには、学習した知識の有効性を定量的に評価し有効であると判定された知識だけを利用する必要がある。これは、一般に知識のユーティリティ (utility of

learned knowledge)を扱う問題と称され、これに関して多くの研究がある[Minton 90, Suwa 92].意味素間の隠喩的結びつきを示すある知識が、どれだけ多くの隠喩文から抽出されたか、また、他の隠喩文の解析に使われ正しい類比を生成することに貢献したか否かなどが、隠喩文理解の問題領域において学習した知識の有効性を測定するための尺度になり得る。

7. 結 論

1階の意味素が一致する1階命題間の類比を拘束条件にして意味記述全体の類比を決定し、異なる意味素間の比喩的關係に関する知識を隠喩文から抽出する方法(プリミティブマッチング法)を提案した。

プリミティブマッチング法は、基本的に異なる意味素間の類比写像を許していること、そして、それにより予想される計算コストの増大を、1階意味素のマッチングという意味的制約を用いることで抑えているこ

との2点に新規性がある。実験を行った結果、二つの意味記述中に共通の意味素が存在する場合に、明らかに類比決定における探索空間を狭める効果があり、意味的制約を併用しない場合に比べ計算量的に有利であることが判明した。また、プリミティブマッチングプロセスによる探索空間の低減率が同程度であるような二つの隠喩文に関してPM法によるコストゲインを比較したところ、使われている動詞の意味記述量が大きいほど、コストゲインが大きいという傾向を得た。この性質は、計算機による言語理解システムとしての本手法の有効性を示唆するものである。

本稿では、特に物理的概念と心的概念の關係に着目し、17個の動詞隠喩文から、三つの心的概念、'attention', 'think', 'know'の物理的イメージに関する知識を抽出し、例証した。プリミティブマッチング法による学習手法は、我々が日常的に使っている言語に潜む「概念間の類似性」に関する知識を顕在化するための基盤技術として有効である。

◇ 参 考 文 献 ◇

- [COBUILD 87] The English Department at the University of Birmingham (ed.): *Collins COBUILD English Language Dictionary*, Collins Pub. (1987).
- [Barnden 89] Barnden, J. A.: Belief, metaphorically speaking, *Proc. 1st Int. Conf. on Principles of Knowledge Representation and Reasoning*, pp. 21-32 (1989).
- [Carbonell 80] Carbonell, J. G.: Metaphor—A key to extensible semantic analysis, *Proc. 3rd Annual Conf. of the Cognitive Science Society*, pp. 292-295 (1980).
- [Falkenheiner 89] Falkenheiner, B., Forbus, K. D. and Gentner, D.: The Structure-Mapping Engine: Algorithm and Examples, *Artif. Intell.*, Vol. 41, pp. 1-63 (1989).
- [Gentner 83] Gentner, D.: Structure Mapping: A theoretical framework for analogy, *Cognitive Science*, Vol. 7, No. 2, pp. 155-170 (1983).
- [Gentner 88] Gentner, D.: Analogical Inference and Analogical Access, Armand Frieditis (ed.), *Analogica*, Morgan Kaufmann (1988).
- [Gentner 89] Gentner, D.: The mechanisms of analogical learning, Vosniadou, S. and Ortony, A. (eds.), *Similarity and analogical reasoning*, chapt. 7, pp. 199-241, Cambridge University Press (1989).
- [Holyoak 89] Holyoak, K. J. and Thagard, P.: Analogical Mapping by Constraint Satisfaction, *Cognitive Science*, Vol. 13, pp. 295-355 (1989).
- [楠見 89] 楠見 孝: メタファの認知モデル, 数理科学, Vol. 307, pp. 39-42 (1989).
- [Lakoff 80] Lakoff, G. and Johnson, M.: *Metaphors we live by*, University of Chicago Press (1980).
- [Martin 90] Martin, J. H.: *A Computational Model of Metaphor Interpretation*, Academic Press (Perspectives in Artificial Intelligence series) (1990).
- [Minton 90] Minton, S.: Quantitative Results Concerning the Utility of Explanation-Based Learning, *Artif. Intell.*, Vol. 42, pp. 363-391 (1990).
- [Norvig 86] Norvig, P.: A unified theory of inference for text understanding, PhD thesis, University of California, Berkeley, CA (1986).
- [Suwa 91] Suwa, M. and Motoda, H.: Learning metaphorical relationships between concepts based on semantic representation using abstract primitives, *Proc. IJCAI 1991 Workshop: Computational Approaches to Non-Literal Language*, pp. 123-131 (1991).
- [Suwa 92] Suwa, M. and Motoda, H.: Learning perceptually-chunked macro-operators, *Proc. Int. Workshop on Machine Intelligence* (1992). (to appear in Machine Intelligence volume 13, Oxford University Press).
- [諏訪 93] 諏訪正樹, 岩山 真: 比喩の計算モデル, 情報処理, Vol. 34, No. 5 (1993).
- [内海 93a] 内海 彰, 堀 浩一, 大須賀節雄: 自然言語処理のための形容詞の意味表現, 人工知能学会誌, Vol. 8, No. 2, pp. 192-200 (1993).
- [内海 93b] 内海 彰, 堀 浩一, 大須賀節雄: 比喩を含む言語表現の理解—形容詞・名詞句と隠喩—, 人工知能学会誌, Vol. 8, No. 2, pp. 201-211 (1993).

[担当編集委員・査読者: 西田豊明]

— 著 者 紹 介 —

諏訪 正樹(正会員)は、前掲(Vol. 9, No. 2, p. 200)参照。

元田 浩(正会員)は、前掲(Vol. 9, No. 1, p. 16)参照。