

## 文脈の中の隠喩 — 関連性による隠喩理解モデル —

内海 彰 菅野 道夫

東京工業大学 総合理工学研究科 システム科学専攻

utsumi@sys.titech.ac.jp

本稿では「XはYだ」という形式の属性隠喩に的を絞り、文脈に依存した隠喩解釈の計算モデルを提案する。特に、文脈情報によって source 概念のどの属性が写像されやすくなるか、という問題を中心に定式化する。文脈に依存した解釈を可能にするために、関連性理論を基盤とした言語解釈の計算モデルを隠喩解釈に適用する。関連性理論に基づく枠組みにおいては、関連性を最大にする解釈がその文脈において最も適切な解釈であるとする関連性の原理に従って言語解釈が行われる。これによって文脈に依存した隠喩解釈が特別な手続きを用いることなく言語解釈の一般的な過程の中で可能になることを示す。

## Contextualized Metaphors — A Relevance Based Model of Metaphor Interpretation —

Akira Utsumi and Michio Sugeno

Department of Systems Science

Interdisciplinary Graduate School of Science and Technology

Tokyo Institute of Technology

utsumi@sys.titech.ac.jp

The purpose of this paper is to propose a computational model of comprehending metaphors in context. This paper is devoted to the role of context in deciding which features of the source concept is preferably mapped to the target concept. A notion of relevance is used for a measure of transferability of features in contextualized metaphors which corresponds to a measure of salience in isolated metaphors. A process of interpreting contextualized metaphors is reasonably integrated with the relevance based framework for interpreting utterances in which the most appropriate interpretation results from searching for an optimal relevance. This suggests that metaphor interpretation needs no specific procedures.

## 1 はじめに

計算機を用いて隠喩の理解過程をモデル化する試みは数多く行われてきたが (e.g., Fass 1991), それらのうち多くの研究は文脈と独立した単一の隠喩文 (isolated metaphor) の理解に費やされ、隠喩理解における文脈の役割はあまり注目されていない。しかし、以下の隠喩文を考えてみると、文脈情報が隠喩の解釈に影響を及ぼすことは明らかである。

### (1) Mary is graceful. John is an elephant.

この例では前文で“graceful”という属性が言及されている。これに続く(1)の下線部の隠喩文においては、概念「象」の持つ“slow”, “heavy”などの様々な属性の中から“graceful”と対照的な“clumsy”が関連性をもつことが期待され、結果として“John is clumsy”という解釈が得られる。しかし一般に“slow”や“heavy”に比べて“clumsy”は象の顕現的な属性ではないため、このような解釈は従来の文脈抜き隠喩理解の枠組みで得ることはできない。このような解釈を必要とする隠喩に対するアプローチとしては、アブダクションによる言語解釈を用いるアプローチ (Hobbs 1991) や談話の結束性に着目したアプローチ (森 & 中川 1992) などが提案されているが、いずれも顕現性 (salience) による従来の隠喩理解の手法 (Iwayama, Tokunaga, & Tanaka 1990) との関係が明確にされていないなど包括的な枠組みまでは明確に示されていない。そこで本稿では「XはYである」という形式の属性隠喩に的を絞り、このような文脈のなかの隠喩 (contextualized metaphor) の理解過程の計算モデルを提案することを目的とする。さらに、提案する計算モデルは隠喩理解に特有ではなく、一般的な言語解釈過程の中で隠喩解釈も暗黙に実現されるようなものを目指す。

一般に、属性隠喩の理解過程は (a) source 概念における顕現的な属性の選択、(b) source 概念の顕現的な属性の target 概念への写像、の2つのステップから成り立つ。前者の属性選択においては、Iwayama et al. (1990) により顕現性の確率的定式化が提案されており、後者の属性写像においては、内海, 堀, & 大須賀 (1993) により写像自体の比喩性について論じられている。しかしいずれのステップにおいても文脈の役割については扱われていない。本稿ではこのうち前者の属性選択についての文脈依存性、つまり文脈情報によって source 概念の属性の典型性がどのように変化するか、を定式化する。

このような文脈依存性を扱うために、本研究では Utsumi & Sugeno (1994) が提案している Sperber & Wilson (1986) の関連性を基盤とした言語解釈の計算モデルを隠喩理解に援用する。この計算モデルでは、「発話などのコミュニケーションにおいて伝達される情報が聞

き手にとって最適な (optimal) 関連性をもつ」という関連性の原理 (Sperber & Wilson 1986) を言語解釈を制約する一般的規約であると見なし、数値的に定義された関連性の値を最適にするような解釈がその文脈において最も適切な解釈であると考えられる。言い替えれば、関連性の原理は、無限の背景知識の中からその発話の解釈に必要な知識を選択するためのヒューリスティクスである。よって source 概念の属性選択における文脈の役割を自然に組み込むことが可能である。

関連性に基づく言語解釈モデルに隠喩理解を統合するためには、source 概念より想起される典型的な属性を文脈の中に組み込み、それらの属性を target 概念に写像するための規則が必要である。そこで本研究では概念の典型的事例 (prototypical exemplar) と属性写像規則 (feature mapping rule) を導入する。典型的事例とはある概念のすべての典型的な記述を満たす架空の個体であり、これを用いて概念のもつデフォルトの属性を命題として矛盾なく文脈に組み込むことが可能である。一方、属性写像規則は概念間の属性写像を許す規則であり、これにより属性の写像が含意という形で言語解釈モデルに統合できる。

## 2 関連性に基づく言語解釈の計算モデル

ここでは関連性に基づく言語解釈の計算モデルについて概説する。詳細については (Utsumi & Sugeno 1994) を参照されたい。

### 2.1 関連性理論の思想

関連性に基づく言語解釈モデルで本質的な役割を果たす関連性は、文脈効果 (contextual effect) と処理努力 (processing effort) という二つの要素から構成される (Sperber & Wilson 1986)。文脈効果とは、聞き手が持っている旧情報 (これらが文脈を構成する) と現在の発話によって得られた情報 (本稿ではこれを想定と呼ぶ) から何らかの新情報を得ることであり、文脈含意 (contextual implication)、文脈強化 (contextual strengthening)、矛盾除去 (elimination of contradictory old assumption) の三種類が考えられている。直観的には、文脈効果とは聞き手が現在持っている情報に対してその発話が何らかの強い影響を与えることであると解釈できる。

もうひとつの処理努力は文脈効果を導出する過程に含まれる心的コストと考えることができる。どのような過程にコストがかかるかを考えると、処理努力は情報にアクセスする際のコストと文脈効果を導出する際のコストに大別することが可能である。しかし後者については

どれだけの文脈効果が得られるかに比例するため相対的には無視することが可能である。よって本稿では「処理努力」は情報へのアクセスに課せられるコストのみを指すこととする。よって、ここで用いている「処理努力」という用語は計算機科学など言うところの処理の計算量を意図しているわけではないことに注意されたい。

これらの二つの要素は trade-off の関係にある。ある情報が特定の文脈において関連性を持つのは、その情報がその文脈において何らかの文脈効果を持つときであり、他の条件が同じであれば文脈効果が大きいほど関連性は高くなる。一方、他の条件が同じであれば処理努力が小さいほど関連性が高くなる。そして関連性の最適値は文脈効果に対する処理努力との最適なバランスを取ることによって計算することができる。

## 2.2 関連性の数値的定式化

関連性に基づく言語解釈の計算モデルにおいては、文脈効果と処理努力は命題に対して定義される確信度 (confidence value) と活性化度 (activation value) によって決定される。つまり、文脈効果の度合いはそれに関連する命題の確信度の差として、処理努力の度合いはその時点において文脈を構成する命題の活性化度の平均値として、それぞれ計算される。

命題の確信度は、0 から 1 までの実数値として与えられ、その命題が表している信念の確からしさを表す指標である。確信度が 0 である命題 ( $s \models \sigma$ ) は、 $s \not\models \sigma$  と同値である。なお確信度は主観確率と異なるものであることに注意する必要がある。そして文脈効果の度合いはその文脈効果の生起以前の確信度と生起後の確信度の差として計算することができる。つまり、文脈効果が得られる前の命題  $p$  の確信度を  $a(t-1)$ 、文脈効果が得られた後のその命題の確信度を  $a(t)$  とすると、上述した 3 種類の文脈効果の度合いは  $|a(t) - a(t-1)|$  となる。例えば、文脈含意の場合では「A氏は医者である」という新たに導出された命題の確信度が 0.8 であるとする、その文脈効果は  $|a(t) - a(t-1)| = |0.8 - 0.0| = 0.8$  となる。文脈  $C$  における想定集合  $A$  の文脈効果の度合い  $CE(A, C)$  は、 $A$  によってもたらされるすべての文脈効果の度合いの和となる。

命題の活性化度は、現在の状態におけるその命題の注目度合いを表すと理解できる。つまり活性化度の大きい命題ほど現在の話題に関係が深いことになり、それだけその命題の accessibility が高いことになる。accessibility が高いということはそれだけ処理努力が少なくて済むことになる。つまり活性化度が大きいほど処理努力が少なくなる。従ってこの活性化度を処理努力の尺度と考えることができる。命題の活性化度は、命題間の依存度 (the degree of dependency) をもとに計算される。依存

度は 2 つの命題の要素間にどれだけの依存関係があるかを表す指標であり、命題  $p_i$  の命題  $p_j$  に対する依存度は (2) 式のように定義される。

$$m(p_i, p_j) = \frac{|E(p_i) \cap E(p_j)|}{|E(p_i)|} \quad (2)$$

集合  $E(p_i)$  は基本的に命題  $p_i$  の構成要素の集合を表しているが、実際には集合の要素とならない構成要素も存在する。詳しくは (Utsumi & Sugeno 1994) を参照のこと。この依存度により、離散時間  $t$  における命題  $p_i$  の活性化度  $c_t(p_i)$  は以下のように計算することができる。なお、この時間  $t$  は談話やテキストなどの一連の流れの中の時点を表しており、時間  $t-1$  は現在対象となっている言語表現の直前の言語表現を解析した時点である。

• 命題  $p_i$  が想定の場合

$$c_t(p_i) = \epsilon \frac{\sum_{p_j \in A_{opt}, i \neq j} m(p_i, p_j)}{|A_{opt}| - 1} + (1 - \epsilon) \frac{\sum_{q_j \in C_M} c_t(q_j) m(p_i, q_j)}{\sum_{q_j \in C_M} c_t(q_j)} \quad (3)$$

上式において、 $A_{opt}$  と  $C_M$  はそれぞれ以下に示す理解手順の step 6 に示される最適な関連性の値をもつときの想定と文脈の命題の集合である。また  $\epsilon$  は定数であり、想定集合における依存度が活性化度全体に与える影響の度合いを表す。本稿では 0.25 とする。また  $|A_{opt}| = 1$  の時には  $\epsilon = 0$  として計算する。(3) 式は、解釈の対象となっている言語表現における依存度が大きいほど活性化度が大きく、文脈との依存度が高いほど活性化が高くなることを意味する。さらに文脈状況における活性化の高いインフオンと依存関係があるほど、活性化は高くなることも意味している。

• 命題  $p_i$  が文脈の場合

$$c_t(p_i) = c_{t-1}(p_i) + \{1 - c_{t-1}(p_i)\} D - \alpha c_{t-1}(p_i) \{1 - D\} \quad (4)$$

$$D = \frac{\sum_{q_j \in A} m(p_i, q_j)}{|A|}$$

上式の  $D$  における  $A$  は関連性に基づく言語解釈モデルの step 2 における想定集合  $A_i$  に相当する。 $D$  は現在の解釈である想定集合との依存度を表しており、現在の解釈を構成する命題との依存度が大きいほど、命題  $p_i$  の活性化度が大きくなることを意味している。また、(4) 式の右辺第 3 項は減衰項であり、想定との依存度が小さいほど減衰が大きいことになる。 $\alpha$  は減衰の程度を表す係数であり、 $\alpha = 0$  は減衰がないことに相当する。また  $\alpha = 1$  のときには  $c_t(p_i) = D$  となり、今までの履歴がまったく影響しない場合に相当する。本稿では例題とする文章が短いので  $\alpha = 0$  と仮定する。

また命題  $p_i$  が辞書的知識から引き出された直後には  $c_{t-1}(p_i)$  の値が存在しないので、(4) 式をそのまま適用

することはできない。そこでこの場合には、(5)式により  $c_{t-1}(p_i)$  の値を初期化する。なお  $a(p_i)$  は命題  $p_i$  の確信度である。

$$c_{t-1}(p_i) = a(p_i) \times \frac{\sum_{q_j \in C_{M+1}(t-1)} c_{t-1}(q_j) m(p_i, q_j)}{\sum_{q_j \in C_{M+1}(t-1)} c_{t-1}(q_j)} \quad (5)$$

つまり過去の話題と関連がある命題ほど活性度が高く、その値はその命題の確信度を越えない値となる。

## 2.3 関連性に基づく言語解釈の計算モデル

関連性に基づく言語解釈の計算モデルにおいては、発話(文)  $P(t)$  の聞き手にとっての関連性の値は以下の手順にしたがって算出される。まず表記法として、時点  $i$  ( $i$  回目の反復) における想定集合を  $A_i(t)$ 、 $A_i(t)$  に対して最適な関連性を達成する文脈 (optimal context) を  $C_i(t)$ 、そのときの最適な関連性の値を  $\tau_i(t)$ 、そのときに  $A_i(t)$  と  $C_i(t)$  から文脈効果で得られた想定集合を  $B_i(t)$  とする。さらに聞き手が発話を理解する時に利用することのできる命題の集合を環境 (environment) と呼び、それを  $W_i(t)$  と表す。この集合の部分集合が  $C_i(t)$  となる。なお以下では、特に断らない限り時間  $t$  を省略して  $A_i = A_i(t)$  と略記する。

このモデルではすべての情報は命題として扱われ、含意は制約によって導出される。本稿における「含意」とは、概略的に言うと、 $C_i \cup A_i$  から導出可能であるが  $C_i$  または  $A_i$  のみからは導出不可能である命題で、かつ  $A_1$  から  $A_{i-1}$  までにおいて含意でない命題のことである。そして前節で定義された文脈効果と処理労力を用いて、 $A_i$  のある文脈  $C_i$  における関連性を定義する。

$$\text{関連性} = \sum_{j=1}^i CE(A_j, C_j) \times \left( \frac{\sum_{p_k \in C_i} c_i(p_k)}{|C_i|} \right)^3 \quad (6)$$

右辺第1項は  $A_i$  の文脈効果の程度、第2項は  $C_i$  に含まれる命題の活性度の平均値をそれぞれ表している。よって式(6)は文脈効果が大きいほど関連性は大きくなり、処理労力が小さいほど(活性度の平均値が大きいほど)関連性は大きくなるという上述した性質を表現している。

**step 0**  $i = 1, \tau_0 = 0.0$ 、発話  $P(t)$  の命題的内容を表す命題の集合を  $A_1$ 、初期環境を  $W_0$  とする。

**step 1**  $W_i = W_{i-1} \cup (A_i \text{ に属する命題に含まれる概念に関する辞書の知識から新たに得られる命題の集合})$  とする。

**step 2**  $A_i$  に対して、 $W_i$  におけるすべての命題の活性度を計算する。

**step 3**  $A_i$  に対して関連性の値が最大となるような文脈  $C_i$  を決定する。

**step 4**  $A_{i+1} = \text{Suc}(A_i)$ 、 $W_{i+1} = \text{Suc}(W_i)$ 、 $i = i + 1$  とする。

**step 5** step 1 から step 4 までのステップを  $P(t)$  の関連性が最適値に達するまで有限回繰り返す。

**step 6**  $\tau_i$  のうちで最大の関連性をもつ時点を  $M$  とすると、 $\tau_M$  がその言語表現のその解釈の聞き手にとっての関連性となる。そして想定集合  $A_{\text{opt}} = A_{M+1} = A_M \cup B_M$  に属する各命題について、最適文脈  $C_M$  に対する活性度を計算する。

集合  $W_{M+1}(t) \cup A_{\text{opt}}(t)$  が次の発話  $P(t+1)$  を解釈するための初期環境  $W_0(t+1)$  とする。

このような手順により最終的に求めた想定集合  $A_{\text{opt}}$  がその言語表現の解釈であると定義する。なお上記の step 4 における  $\text{Suc}(A_i)$  と  $\text{Suc}(W_i)$  は概略的に言って、それぞれ  $A_i \cup B_i -$  (消去された  $A_i$  の命題の集合) と (文脈強化により確信度が変化した後の状態の  $W_i -$  (矛盾除去によって消去された命題の集合)) を表している。

また上記の手順において関連性の最適値を求める部分が step 3 と step 5 の2箇所存在する。これらの箇所においてどのように最適値を求めるかが問題であるが、一般に文脈を構成する可能性のある要素は有限ではないので実際に最適解が求まるかどうかは決定できない。また実際に人間が言語表現を理解するときには厳密に最適な値を求めているとは考えられず、状況に応じてどの程度の近似解を求めるかが決定される。そこで本研究では以下のような基準により最適であると考えられる。

- step 3 の文脈の選択においては、環境  $W_i$  のすべての命題をその活性度の値によって降順にソートして、活性度の大きさが  $j$  番目以内の命題の集合を  $S_{ij}$  とする。そして  $j = 1$  から順に文脈を  $S_{ij}$  としたときの関連性の値  $u_{ij}$  を計算し、 $CE(A_i, S_{ij}) > CE(A_i, S_{ij-1})$  を満たす(つまり新しい文脈効果が得られる)文脈  $S_{ij}$  を求める。このような文脈は最適文脈になる可能性があるので潜在文脈 (potential context) と呼ぶ。そして二つの隣接した潜在文脈  $S_{ik}, S_{il}$  ( $k < l$ ) (つまり  $k < m < l$  を満たすすべての  $m$  について  $S_{im}$  は潜在文脈ではない) において、最初に  $u_{ik} > u_{il}$  を満たす文脈  $S_{ik}$  を  $C_i$  とする。もしこのような潜在文脈がない場合には、最も大きい関連性の値をもつ潜在文脈を  $C_i$  とし、潜在文脈自身が存在しない時には  $C_i = \phi$ 、 $\tau_i = 0.0$  とする。
- step 5 においても、最初に  $\tau_i > \tau_{i+1}$  を満たす  $i$  を  $M$  とする。

奇麗さ	0.90	{ 汚い: 0.98 きれい: 0.02 }
整頓さ	0.70	{ 乱雑な: 1.00 }
におい	0.60	{ くさい: 1.00 }
明るさ	0.50	{ 暗い: 0.80 明るい: 0.20 }
広さ	0.40	{ 狭い: 1.00 }

図 1: 概念「豚小屋」の確率的表現

### 3 文脈の中の隠喩の理解

まずは次の隠喩文を例にして考えてみよう。

(7) この部屋は豚小屋だ。

この隠喩文だけが単独で与えられたとすると、概念「豚小屋」の典型的な属性である「汚い」、「乱雑な」や「くさい」などが概念「この部屋」に写像され、結果として「この部屋は汚くて、乱雑で、くさい」というような解釈が得られるであろう。このような文脈のない隠喩文の解釈においては、source 概念を構成する多くの属性のうちどの属性が高い顕現性を持つかを決定することが重要になる。

Iwayama et al. (1990) は概念を構成する属性の顕現性を数値的に計算する手法を提案し、この顕現性の基準によりどの属性が優先的に写像されるかを処理している。source 概念や target 概念はプロトタイプ理論におけるプロトタイプに確率値を割り当てた図 1 のような表現方法が用いられている。例えば、この図において「奇麗さ」や「広さ」などが属性名であり、その右側の実数値がその性質の重要度  $d_i$  を表している。さらに右側の { } で囲まれた部分が属性値集合  $V_i$  に相当する。属性値集合は「汚い: 0.98」のように属性値とその確率値の対  $v_{ij} : p_{ij}$  の集合である。属性値の確率は概念の性質を思い浮かべるときにその属性値がその概念の典型的な属性値であると判断する確率と考えることができる。そして各属性値集合の中で確率値が最も大きい属性値のことを、その属性の最尤属性値と呼ぶ。例えば、属性「奇麗さ」においては最尤属性値は「汚い」である。

このような概念表現に対して、Iwayama et al. は性質  $S_i$  の顕現性の値の計算式として以下の式を用いた。

$$\text{顕現性}(S_i) = d_i \times (1.0 - H_i) \quad (8)$$

$$H_i = \begin{cases} \frac{\sum_{j=1}^m p_{ij} \log_2 \frac{1}{p_{ij}}}{\log_2 m} & (m \geq 2) \\ 0 & (m = 1) \end{cases}$$

この式において  $m = |V_i|$  である。例えば、概念「豚小屋」における性質「奇麗さ」の顕現性は、

$$0.90 \times (1.0 - 0.98 \log_2 \frac{1}{0.98} - 0.02 \log_2 \frac{1}{0.02}) = 0.77$$

と計算することができる。同様にして、図 1 における各性質の顕現性は以下ようになる。

奇麗さ: 汚い	$0.90 \cdot 0.86 = 0.77$
整頓さ: 乱雑な	$0.70 \cdot 1.0 = 0.70$
におい: くさい	$0.60 \cdot 1.0 = 0.60$
広さ: 狭い	$0.40 \cdot 1.0 = 0.40$
明るさ: 暗い	$0.50 \cdot 0.28 = 0.14$

Iwayama et al. の手法では、ある閾値を設定して顕現性がこれより大きい値をもつ属性が写像される。例えば (7) の解釈において閾値を 0.5 と設定すると、上記の結果より「この部屋は汚くて、乱雑で、くさい」という解釈が得られることになる。

しかしこの隠喩文がある特定の文脈下に置かれると、単純に概念の顕現性を計算するだけでは適切な解釈が得られない。

- (9) a. 私は狭い部屋には絶対に住みたくない。
- b. この部屋は豚小屋だ。

この文章では隠喩文 (9b) は (7) と同一の文であるが、その解釈は異なるのが自然である。つまり直前の文 (9a) が顕現性が低い属性である「広さ: 狭い」と関連するため属性「狭い」が優先的に写像されることになり、結果として「この部屋は狭い」という異なった解釈が得られる。このような文脈による属性選択への影響は、Gildea & Glucksberg (1983) によって心理学的にも実証されている。よって以下では、本稿の手法の有効性を示すために本手法が文脈情報を考慮しない場合と考慮する場合とで同一の隠喩文 (7) の解釈が異なるという現象を説明できることを示す。

まずは文脈の中の隠喩文である (9b) の解釈過程を見てみよう。図 2 は関連性に基づく枠組みにより文 (9a) が解釈された結果から得られた隠喩文 (9b) を解釈するための初期文脈  $W_0$  を示している。この図において、 $(s_{des} \models \ll \Rightarrow, \ll \text{探る}, \dot{x}, \dot{y} \gg \wedge \ll \text{部屋}, \dot{y} \gg, s_{desire: \dot{x}} \models \ll \text{住む}, \dot{x}, \dot{y} \gg \gg)$  は「もし X が住むための部屋 Y を見つければ、そこに住みたい」という一般的な制約を意味し、文 (9a) の解釈にあたって話し手がある部屋を見つけたと仮定してその命題から弱い含意を導き出している。なお  $s_{desire: 私}$  は「私」の望んでいることの状態 (desire situation) を表している。

このような状況において、隠喩文 (9b) の命題的内容である  $(s_{des} \models \ll \text{豚小屋}, x \gg)$  が  $A_1$  を形成する。そし

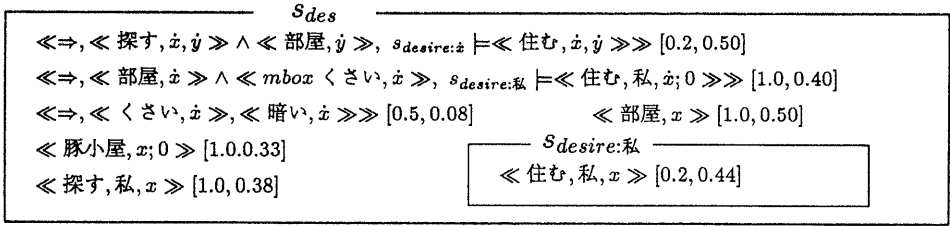


図 2: 隠喩文 (9b) の解釈のための初期環境

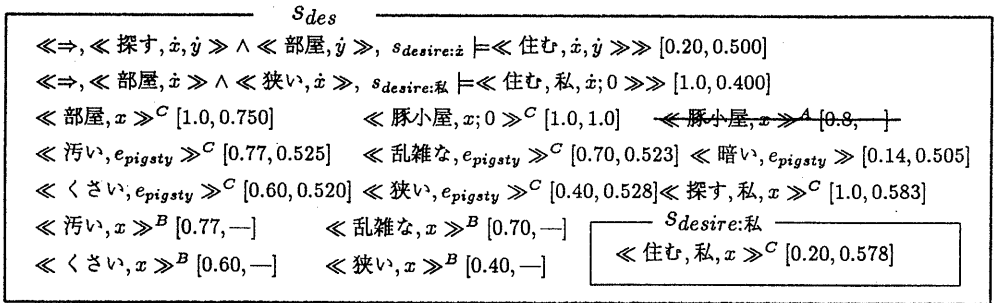


図 3: 隠喩文 (9b) の解釈時点 1 における環境

て step 1 において概念「豚小屋」から得られる辞書的知識として、この概念の典型的な属性が命題の形で  $W_1$  に加えられる。

$$\begin{aligned}
 (s_{des} \models \langle \text{汚い}, e_{pigsty} \rangle) [0.77, 0.525] \\
 (s_{des} \models \langle \text{乱雑な}, e_{pigsty} \rangle) [0.70, 0.523] \\
 (s_{des} \models \langle \text{小さい}, e_{pigsty} \rangle) [0.60, 0.520] \\
 (s_{des} \models \langle \text{狭い}, e_{pigsty} \rangle) [0.40, 0.529] \\
 (s_{des} \models \langle \text{暗い}, e_{pigsty} \rangle) [0.14, 0.505]
 \end{aligned} \tag{10}$$

これらの命題に使われている  $e_{pigsty}$  が概念「豚小屋」の典型的事例を表す個体である。典型的事例はその概念の最も典型的な個体であり、その概念に対して人間が持っている一般的な知識をすべて満たす。このような架空の個体を用いる理由としては、その概念に属するすべての個体がそれらの典型的な属性を備えているわけではないということのほか、一般的なデフォルトとしての知識と具体的な個体に関する知識の区別を明確にしたいという理由もある。(10)において右側の括弧内の二つの数値はその命題の確信度と活性度を表している。それぞれの確信度は前述した顕現性の値をそのまま用いている。これは顕現性の高い属性ほどその概念にその属性が備わっていると感じる度合いが高いという自然な仮定である。活性度については、step 2 で式 (5) で初期化したあとに、式 (4) で計算した値を示してある。式 (5) の定義により、文脈と独立した隠喩文を理解する場

合には顕現性が高い属性ほど活性度も高くなる。

このようにして step 2 で活性度の計算された  $W_1$  は活性度の大きい順にソートされる。ここで注目すべきことは (10) からわかるように概念「豚小屋」から得られた命題のうち、「狭い」という属性の活性度が一番大きいということである。つまり顕現性の高い属性である「汚い」や「乱雑な」といった知識よりも accessibility が高い、つまり target 概念に写像される可能性が高いということを表している。そして step 3 ではこれらの命題により局所的な文脈が増進的に生成され、それぞれの文脈における関連性が計算される。例えば最初は活性度の一番大きい命題である  $(s_{des} \models \langle \text{豚小屋}, x; 0 \rangle)$  のみからなる文脈  $S_{11}$  が構成され、 $A_1$  に対して文脈効果があるかどうか調べられる。この場合には隠喩文 (9b) から直接得られる命題と  $S_{11}$  の命題が矛盾し、確信度の小さい前者が消去されるために文脈効果 (矛盾除去) にはならない。よって  $S_{11}$  は潜在文脈ではなく、次に活性度が二番目に大きい命題である  $(s_{des} \models \langle \text{部屋}, x \rangle)$  が加えられた文脈  $S_{12}$  がチェックされていく。

この例において最初の潜在文脈は 5 個の命題からなる文脈  $S_{15}$  である。この文脈に含まれている辞書的知識から得られた命題である  $(s_{des} \models \langle \text{狭い}, e_{pigsty} \rangle)$  が属性写像規則により target 概念である「この部屋」に

<i>S<sub>des</sub></i>	
$\langle\langle \Rightarrow, \langle\langle \text{探す}, \dot{x}, \dot{y} \rangle\rangle \wedge \langle\langle \text{部屋}, \dot{y} \rangle\rangle, s_{desire:\dot{x}} \models \langle\langle \text{住む}, \dot{x}, \dot{y} \rangle\rangle \rangle\rangle^C$	[0.20, 0.500]
$\langle\langle \Rightarrow, \langle\langle \text{部屋}, \dot{x} \rangle\rangle \wedge \langle\langle \text{狭い}, \dot{x} \rangle\rangle, s_{desire:私} \models \langle\langle \text{住む}, 私, \dot{x}; 0 \rangle\rangle \rangle\rangle^C$	[1.0, 0.430]
$\langle\langle \Rightarrow, \langle\langle \text{くさい}, \dot{x} \rangle\rangle, \langle\langle \text{暗い}, \dot{x} \rangle\rangle \rangle\rangle$	[0.50, 0.063]
$\langle\langle \Rightarrow, \langle\langle \text{部屋}, x \rangle\rangle^C$	[1.0, 0.750]
$\langle\langle \Rightarrow, \langle\langle \text{豚小屋}, x; 0 \rangle\rangle^C$	[1.0, 0.667]
$\langle\langle \Rightarrow, \langle\langle \text{汚い}, e_{pigsty} \rangle\rangle$	[0.77, 0.169]
$\langle\langle \Rightarrow, \langle\langle \text{乱雑な}, e_{pigsty} \rangle\rangle$	[0.70, 0.165]
$\langle\langle \Rightarrow, \langle\langle \text{暗い}, e_{pigsty} \rangle\rangle$	[0.14, 0.009]
$\langle\langle \Rightarrow, \langle\langle \text{くさい}, e_{pigsty} \rangle\rangle$	[0.60, 0.159]
$\langle\langle \Rightarrow, \langle\langle \text{狭い}, e_{pigsty} \rangle\rangle$	[0.40, 0.175]
$\langle\langle \Rightarrow, \langle\langle \text{探す}, 私, x \rangle\rangle^C$	[1.0, 0.583]
$\langle\langle \Rightarrow, \langle\langle \text{汚い}, x \rangle\rangle^B$	[0.77, 0.367]
$\langle\langle \Rightarrow, \langle\langle \text{乱雑な}, x \rangle\rangle^B$	[0.70, 0.367]
$\langle\langle \Rightarrow, \langle\langle \text{くさい}, x \rangle\rangle^B$	[0.60, 0.367]
$\langle\langle \Rightarrow, \langle\langle \text{狭い}, x \rangle\rangle^B$	[0.40, 0.440]
<i>S<sub>desire:私</sub></i>	
$\langle\langle \Rightarrow, \langle\langle \text{住む}, 私, x; 0 \rangle\rangle^B$	[0.40, 0.357]

図 4: 隠喩文 (9b) の解釈における最適な文脈

<i>S<sub>des</sub></i>	
$\langle\langle \Rightarrow, \langle\langle \text{探す}, \dot{x}, \dot{y} \rangle\rangle \wedge \langle\langle \text{部屋}, \dot{y} \rangle\rangle, s_{desire:\dot{x}} \models \langle\langle \text{住む}, \dot{x}, \dot{y} \rangle\rangle \rangle\rangle$	[0.20, 0.00]
$\langle\langle \Rightarrow, \langle\langle \text{豚小屋}, x; 0 \rangle\rangle^C$	[1.00, 1.00]
$\langle\langle \Rightarrow, \langle\langle \text{汚い}, e_{pigsty} \rangle\rangle^C$	[0.77, 0.385]
$\langle\langle \Rightarrow, \langle\langle \text{乱雑な}, e_{pigsty} \rangle\rangle^C$	[0.70, 0.350]
$\langle\langle \Rightarrow, \langle\langle \text{くさい}, e_{pigsty} \rangle\rangle^C$	[0.60, 0.300]
$\langle\langle \Rightarrow, \langle\langle \text{狭い}, e_{pigsty} \rangle\rangle$	[0.40, 0.200]
$\langle\langle \Rightarrow, \langle\langle \text{暗い}, e_{pigsty} \rangle\rangle$	[0.14, 0.070]
$\langle\langle \Rightarrow, \langle\langle \text{汚い}, x \rangle\rangle^B$	[0.77, 0.403]
$\langle\langle \Rightarrow, \langle\langle \text{乱雑な}, x \rangle\rangle^B$	[0.70, 0.399]
$\langle\langle \Rightarrow, \langle\langle \text{くさい}, x \rangle\rangle^B$	[0.60, 0.391]

図 5: 文脈と独立した隠喩文 (7) の解釈における最適な文脈

写像される。属性写像規則は一般に以下の形式を取る。

$$\langle\langle R, a(es) \rangle\rangle \wedge \langle\langle T, \dot{x} \rangle\rangle \wedge \langle\langle DC, \langle\langle S, \dot{x} \rangle\rangle \rangle \Rightarrow \langle\langle F(R, S, T), a(\dot{x}) \rangle\rangle \quad (11)$$

式 (11) において  $es$  は概念  $S$  の典型的事例,  $T$  は target 概念,  $a(x)$  は引数の並びをそれぞれ表している。またインフオン  $\langle\langle DC, s \models \langle\langle S, \dot{x} \rangle\rangle \rangle$  は発話の意味 (記述的内容) が  $\langle\langle S, \dot{x} \rangle\rangle$  であることを示している。この規則の意味は「概念  $S$  に情緒的属性  $R$  があるとき, その属性を別の概念  $T$  に写像すると, 「概念  $T$  が属性  $F(R, S, T)$  をもつ」という (文脈) 含意が得られる」となる。なお単純には属性  $R$  が概念  $T$  でも成り立つ場合には  $F(R, S, T) = R$  となるが, 一般的には属性写像自体が比喩的になる場合が多く, 必ずしもこのようになるとは限らない。このような場合についての  $F(R, S, T)$  の計算方法については, 内海ほか (1993) により提案されているが, ここではとりあえず考えないことにする。またこのような属性写像規則は隠喩解釈のための特別な規則ではなく, 字義どおりの解釈をもつ通常の言語解釈一般にも必要な規則であることに注意されたい。例えば「彼女はハリウッドの女優であるが, あまり美しくない」という文を考えてみよう。前半節と後半節が逆接を表す接続助詞「が」で結ばれているので, 前半節より導かれる「彼女は美しい」という命題を後半節が否定していると解釈するのが最も妥当である。つまり概念「ハリウッドの女優」の典型的な属性である「美しい」が「彼女」に写像されてい

るということになり, このことから比喩的ではない通常の文においてもこのような属性の写像が行われていることがわかる。よって写像規則 (11) は字義的/非字義的に関係なく言語解釈一般に必要な規則であるということが出来る。 $S_{15}$  においては (11) による属性「狭い」の写像は以下ようになる。

$$\langle\langle \text{狭い}, e_{pigsty} \rangle\rangle \wedge \langle\langle \text{部屋}, x \rangle\rangle \wedge \langle\langle DC, s_{des} \models \langle\langle \text{豚小屋}, x \rangle\rangle \rangle \Rightarrow \langle\langle \text{狭い}, x \rangle\rangle$$

含意された命題 ( $s_{des} \models \langle\langle \text{狭い}, x \rangle\rangle$ ) の確信度は ( $s_{des} \models \langle\langle \text{狭い}, e_{pigsty} \rangle\rangle$ ) と同じ 0.4 なので, (6) 式により文脈  $S_{15}$  における関連性の値  $u_{15}$  は

$$u_{15} = 0.40 \times \left( \frac{1.000 + 0.750 + 0.583 + 0.578 + 0.529}{5} \right)^3 = 0.130$$

と求められる。同様に  $S_{16}$  以降も新たな文脈効果が得られていくが,  $u_{16} = 0.338, u_{17} = 0.493, u_{18} = 0.606, u_{19} = 0.600$  となり,  $S_{18}$  が  $u_{ik} > u_{il} (k < l)$  を満たす最初 (最小) の潜在文脈となる。よって  $C_1 = S_{18}, r_1 = u_{18} = 0.606$  となる。これらの結果を示したのが図 3 である。この図において  $C$  の付いている命題は最適文脈  $C_1$  の要素,  $B$  の付いている命題 (つまり含意) は  $B_1$  の要素であることをそれぞれ示している。

これらのプロセスは最適な関連性が達成されるまで続けられる。新しい想定集合である  $A_2 = Suc(A_1) = A_1 \cup B_1 - \{ (s_{des} \models \langle\langle \text{豚小屋}, x \rangle\rangle) \}$  に対する関連性の最適値が同じようにして計算される。 $A_2$  に対し

て最適な関連性を達成する文脈は  $S_{26}$  であり、ここでは  $A_2$  の要素である ( $s_{des} \models \ll \text{狭い}, x \gg$ ) と (9a) の解釈である制約 ( $s_{des} \models \ll \Rightarrow, \ll \text{部屋}, x \gg \wedge \ll \text{狭い}, x \gg, s_{desire:私} \models \ll \text{住む}, 私, x; 0 \gg$ ) から ( $s_{desire:私} \models \ll \text{住む}, 私, x; 0 \gg$ ) (私はこの部屋には住みたくない) という含意が得られる。この含意は旧情報である ( $s_{desire:私} \models \ll \text{住む}, 私, x \gg$ ) と矛盾するが、こちらの確信度のほうが小さいため矛盾除去というさらなる文脈効果が得られる。関連性の値  $r_2 = u_{26} = 0.614$  は  $r_1$  より大きいためこの過程はさらに続くことになるが、この例題においてはこれ以上関連性が最適にはならないので、最終的に  $C_2 = S_{26}$  が  $C_M$  となる。この最終的な結果を示したのが図 4 である。結果として関連性に基づく言語解釈の枠組みにより、文脈の中の隠喩文 (9b) の解釈として「この部屋は狭くて、汚くて、乱雑で、くさくて、「私」はこの部屋には住みたくないと思っている」という、直観的に妥当な解釈が得られる。

次に文脈に依存した隠喩の解釈と文脈から独立した隠喩の解釈の違いを明らかにし、本稿の手法がいずれの隠喩も適切に解釈できることを示すために、(9b) と同じ隠喩文で、かつ (9a) のような関連した文脈のない (7) の解釈結果を図 5 に示しておく。文脈と独立したこの隠喩の解釈過程においては、辞書知識から得られる命題の活性度の定義式 (5) から明らかのように顕現性の値と活性度の値が比例するため、概念「豚小屋」の典型的な属性が顕現性の高い順にサーチされる結果、(Iwayama et al. 1990) の手法と同じ順番で属性が写像されていく。そして属性「狭い」の (含意としての) 写像ではより高い関連性を達成することができないため、図 5 のように (7) の解釈は「この部屋は汚くて、乱雑で、くさい」となる。

#### 4 おわりに

本稿では関連性に基づく言語解釈過程に隠喩理解を適用し、文脈に依存した隠喩解釈が可能であることを示した。この解釈モデルにおいては関連性という概念が、(a) 現在の文脈に関連する source 概念の属性を優先的に target 概念に写像させ、他の属性の写像を抑制する (b) 属性の写像を含む実際に導出可能な非常に多くの含意のうち、現在の文脈において特に関連の高い含意の導出を選択する、という 2 つの重要な役割を果たしている。さらに、本稿では述べなかったが、本モデルは字義的な文と隠喩などの非字義的な文の解釈過程には本質的な違いはないという最近の心理学の知見に沿ったものである。なお本稿で示したモデルや例題は prolog プログラムとして計算機に実装されている。

しかしながら本稿の手法はまだ不十分な点も多い。

Hobbs (1991) のモデルにおいては (1) に接続詞を挿入した “Mary is graceful, but John is an elephant.” のような文における接続詞の役割に注目している。本稿のモデルは現時点ではこのような結束性を扱うことはできない。しかしテキストにおける接続詞の役割を関連性の概念を用いて説明する研究は Blakemore (1992) などによってなされており、本稿のモデルにもこのような結束性を扱う能力は十分に備わっていると考えられる。また、対象とする隠喩は今のところ属性隠喩に限られているが、構造隠喩にまで範囲を広げることも必要である。

#### 参考文献

- Blakemore, D. (1992). *Understanding Utterances*. Oxford: Basil Blackwell.
- Fass, D. (1991). met\*:a method for discriminating metonymy and metaphor by computer. *Computational Linguistics*, 17(1), 49–90.
- Gildea, P. & Glucksberg, S. (1983). On understanding metaphor: The role of context. *Journal of Verbal Learning and Verbal Behavior*, 22, 577–590.
- Hobbs, J. (1991). Metaphor and abduction. In Fass, D., Hinkelman, E., & Martin, J. (Eds.), *Proceedings of the IJCAI Workshop on Computational Approaches to Non-Literal Language: Metaphor, Metonymy, Idioms, Speech Acts, Implicature*, pp. 52–61.
- Iwayama, M., Tokunaga, T., & Tanaka, H. (1990). A method of calculating the measure of salience in understanding metaphors. In *Proceedings of the Eighth National Conference on Artificial Intelligence (AAAI-90)*, pp. 298–303.
- 森 辰則, 中川 裕志 (1992). 談話の結束性と比喩理解. 日本認知科学会第 9 回大会発表論文集, pp. 50–51.
- Sperber, D. & Wilson, D. (1986). *Relevance: Communication and Cognition*. Oxford, Basil Blackwell. 内田 聖二, 中遠 俊明, 宋 南先, 田中 圭子 (訳) (1993). 関連性理論 — 伝達と認知 —, 研究社出版.
- 内海 彰, 堀 浩一, 大須賀 節雄 (1993). 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, 8(2), 201–211.
- Utsumi, A. & Sugeno, M. (1994). Implicature and relevance: A computational model of relevance theory. *submitted for publication*.