

概念ベースにおける動詞の精練

田中 翔[†] 渡部 広一[‡] 河岡 司[‡]

[†]同志社大学大学院工学研究科 〒610-0394 京都府京田辺市多々羅都谷 1-3

E-mail: [†] dtg0741@mail4.doshisha.ac.jp, [‡] {hwatabe, tkawaoka}@mail.doshisha.ac.jp

あらまし 本研究では、自然言語処理における常識的な判断を目的とする。そして、常識的な判断を実現するために、連想メカニズムを利用している。連想メカニズムは、電子辞書等より機械的に構築した概念ベースと概念間の強さを定量化する関連度計算により構成される。概念ベースの概念を見ると、動詞概念は正しく特徴付けられていない概念が多い。しかし、動詞は連想において重要な役割を果たすので、正しく特徴付ける必要がある。本稿では動詞概念に着目し、動詞と名詞の関係を体系化したデータベースである Web から自動構築した大規模格フレームを用いて動詞概念を追加、精練する手法と概念ベースを用いて動詞概念を追加、精練する手法の2つを提案する。

キーワード 概念ベース、関連度、連想メカニズム、動詞

Refinement of Verb in Concept-Base

Sho TANAKA[†] Hirokazu WATABE[‡] and Tsukasa KAWAOKA[‡]

[†] Graduate School of Engineering, Doshisha University

1-3 Miyakodani Tatara Kyotanabe-shi, Kyoto, 610-0394 Japan

E-mail: [†] dtg0741@mail4.doshisha.ac.jp, [‡] {hwatabe, tkawaoka}@mail.doshisha.ac.jp

Abstract In this study, in the natural language processing, it aims to judge based commonsense. And to realize a commonsense judgment it is necessary to the association mechanism. The association mechanism is composed by Concept-base which is constructed by analyzing the Electronic Dictionary etc. and the calculation method of the degree of association to quantify strength between concepts. When the Concept-base is analyzed about concept, verb concept is not correctly characterized. As verb concept plays an important role in association, it is necessary to characterize verb concept correctly. In this paper, it pays attention to the verb, and it proposes two of the techniques for the addition, and refining the attribute. First, the technique is using Case Frame Compilation from the Web using High-Performance Computing, the database systematized relation between the verb and the noun. Second, it is a technique for the addition, and refining the attribute of the verb concept with the Concept-base.

Keyword Concept-Base, Degree of Association, Association Mechanism, Verb

1. はじめに

近年、情報処理技術は急速に発展してきており、人間の生活になくてはならない技術となっている。本研究では、情報処理技術の応用として着目されている自然言語処理において、常識的な判断を行うことを目的としている。常識的な判断とは、論理判断、量的判断、時間判断、場所判断、感覚判断、感情判断を指す。これら常識的な判断を実現するのが常識判断メカニズムである。常識判断メカニズム^[1]において、全ての知識を定義することは不可能であるため、知識としてないものに対しては、連想を行い、知識を補う必要がある。その連想処理を実現するために、概念を属性と重みの対の組で定義した概念ベースと、概念間の関連の強さを定量的に評価するための関連度計算方式などで構成する連想メカニズムを利用する。図1に常識判断メカニズムと連想メカニズムの全体像を示す。

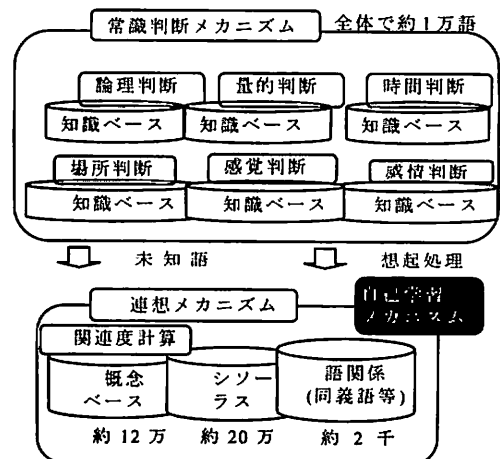


図1. メカニズム全体図

連想メカニズムとは図1に示したように概念ベース^[2]と関連度計算方式^{[3], [4]}, 語彙体系であるシソーラス^[5]や同義語・類義語・反対語といった語関係データベースによって構成されている。また, それらを利用し, 想起語処理と未知語処理^[1]を実現することが可能になる。想起語処理とは, 例えば, 「真珠」という言葉はコンピュータにとっては文字列にしか過ぎないが, 人間であれば「綺麗」「高価」「宝石」などといった言葉を想起することができる。このような処理をシステムとしてコンピュータ上で実現するものである。この想起により, コンピュータは「真珠を貰った」という, 入力文に対して, 「高価」なものを貰ったというを理解することができ, 会話へつなげることが可能となる。また, 未知語処理とは, 「ゴールデンレトリバー」などの常識判断メカニズム内の各システム内の知識ベース^[2]に定義されていない言葉(未知語)を, 「犬」という知識ベースに定義されている言葉(既知語)と置き換える処理のことであり, この処理によって, コンピュータに与える知識を最小限に留めることが可能となる。

これらの連想機能に用いられる概念ベースとは, 語(概念)をその概念を特徴付ける語である属性とその重要性を示す重みとの組み合わせによって表現した大規模データベースである。概念ベースの構築方法として, 電子化された辞書を利用して機械的に構築する手法^[2]が提案されている。しかし, 従来の概念ベースは, 動詞概念に対する特徴の語の抽出が困難であるという問題がある。そのため, 動詞に対して適切な連想を行えない。しかし, 会話や文章にはほとんどの場合, 動詞が含まれており, 動詞概念はコンピュータが人間のように柔軟な判断をするために非常に重要な役割を果たす。

本稿では, 従来の現在の概念ベースに動詞の属性の追加, 精練を行い, 動詞からの連想を適切に行うための手法を提案する。

2. 連想メカニズム

それぞれの常識判断システムに入力される多くの語は, それぞれの知識ベースに登録された語ではない。これらに対する常識的な判断も必要であり, すべての語に対する判断が必要である。それぞれの知識ベースに登録されていない語を未知語と呼び, それらの語に対しても常識判断を行うメカニズムが必要となる。

未知語に対して, 知識ベースを利用するためには, 連想メカニズムを構成する概念ベース(2.1節)や関連度計算(2.2節)を用いることにより, これらの未知語について, 意味的な関連やその強さの度合いを評価し, 最も関連の強い代表語を決定する必要がある。この処理を総称して「未知語処理」と呼ぶ。

2.1. 概念ベース

概念ベースとは, 電子化された国語辞書などを機械的に解析し, 辞書の見出し語を概念表記とし, 各見出し語に付与されている説明文を形態素解析することによって得られる自立語をその概念を表す属性とし, さらに, それら属性がその概念にとってどの程度価値のあるものかを表す重みの対の集合として定義した大規模知識ベース^[6]である。概念ベースに定義される語 A に対して L 個の属性が付与されているとき, 属性 a_i を, 重みを w_i とし, 概念 A は次式で表される。

$$\text{概念}A = \{(a_1, w_1), (a_2, w_2), \dots, (a_i, w_i), \dots, (a_L, w_L)\}$$

現在, 概念ベースは国語辞書より機械的に構築したものの^{[2], [7]}に加え, 新聞記事から概念と属性を取得し, 概念ベースに登録することにより約12万語の概念で構成されている。また, 各概念に対し平均30個の属性が付与され, 属性に付与される重みは, 各概念について, その総和が1.0となるように正規化されている。

各概念に付与されている属性は, 概念ベースに概念として定義されている語で構成されるため, 各属性を一つの概念としてみなした場合, さらに表す属性を導くことができる。このように, 概念は概念ベースにより n 次の属性連鎖集合として定義する。また, n 次の属性集合を n 次属性と呼ぶ。なお, 次に述べる関連度計算方式では, 属性連鎖により概念を2次属性まで展開し演算を行う。以下「1次属性」を単に「属性」とも表記する。

2.2. 関連度

概念ベースを利用した概念間の類似性を定量的に評価する手法として, ベクトル空間モデルが広く用いられている。しかし, 本研究では, 概念を定義する属性集合の重みを含めた一一致度を基本とした関連度計算方式を利用し, 概念間の関連性を評価対象としている。これは, 関連度計算方式が有限ベクトル空間によるベクトル空間モデルよりも良好な結果が得られるという報告がなされているためである。

2.2.1. 一一致度

概念 A, B をその1次属性を a_i, b_j , 重みを u_i, v_j とし, 属性がそれぞれ L 個, M 個($L \leq M$)とすると

$$A = \{(a_1, u_1), (a_2, u_2), \dots, (a_L, u_L)\}$$

$$B = \{(b_1, v_1), (b_2, v_2), \dots, (b_M, v_M)\}$$

と表現する。概念 A, B の一一致度 $MatchWR(A, B)$ は

$$MatchWR(A, B) = \sum_{a_i=b_j} \min(u_i, v_j)$$

$$\min(\alpha, \beta) = \begin{cases} \alpha & (\beta \geq \alpha) \\ \beta & (\alpha > \beta) \end{cases}$$

(各概念の重みの総和は1に正規化する)

と定義する。このとき、一致度は一致する属性のうち小さい方の重みとなるが、これは両方の属性に共通して存在する重み分は有効だと考えるためである。

2.2.2. 関連度

関連度とは、2つの概念間における関係の深さを定量化した0~1の実数値で、関係が深いほど大きな値となる^[8]。概念ベースの属性と重みを再帰的に参照することにより計算される。例として、概念「電車」に対する、「汽車」、「自動車」、「馬」の関連の強さを表1に示す。

表1. 関連度の例

基準概念	対象概念	関連度
電車	汽車	0.785
	自動車	0.135
	馬	0.016

関連度は、対象となる全ての1次属性の組み合わせについて一致度を計算し、1次属性どうしの対応を決定することにより計算する。具体的には、一致する1次属性どうしについては優先的にその対応を決定する。一致しない部分については、その一致度の合計が最大になるように1次属性どうしの対応を決定する。このように対応を決め、対応の取れた属性の組み合わせがT個の場合、概念A、Bの意味関連度 $Rel(A,B)$ は、 $MR(A,B) =$

$$\frac{\sum_{i=1}^T MatchWR(a_i, b_i) \times (u_i + v_i) \times (\min(u_i, v_i) / \max(u_i, v_i))}{2}$$

2.3. 従来概念ベース評価

2.3.1. 評価用データ

概念ベースの評価尺度として、4組1セットとなる概念表記群を準備する。これを(X-A,B,C)データとよび、任意の基準概念Xに対し、関連が深いと考えられる概念A、関連があると思われる概念B、そして全く無関係である概念Cによって構成する。このような評価用データを作成し、1780組のデータを用意した。また、本稿では、動詞概念の精練であるので、1780組のデータの中から、基準概念X、概念A、概念Bに動詞を用いているデータを抜き出し、動詞評価用のデータとして314組も用意した。

2.3.2. 評価方法

概念Xと概念Aとの関連度 $Rel(X,A)$ 、概念Xと概念Bとの関連度 $Rel(X,B)$ 、概念Xと概念Cとの関連度 $Rel(X,C)$ が、

$$Rel(X,A) - Rel(X,B) > Ave(Rel(X,C))$$

$$Rel(X,B) - Rel(X,C) > Ave(Rel(X,C))$$

$$Ave(Rel(X,C)) = \frac{\sum_{i=1}^m Rel(X_i, C_i)}{m}$$

を満たすときを正解とする。mは評価セット数である。このとき、評価用データ全体に対し、正解したものの比率をC平均順序正解率とする。関連度計算方式で、評価用データでのC平均順序正解率が高ければ、人間の感覚に近い判断ができているといえる。また、C平均順序正解率が高ければ、未知語処理においては、未知語を適切な既知語に置き換えることが可能となり、想起語処理では、関連の深い語の順に想起語を出力することが可能となる。

2.3.3. 評価

従来概念ベースのC平均順序正解率の評価結果を表2に示す。

表2. 従来概念ベースのC平均順序正解率

	1780組	動詞評価用
C平均順序正解率(%)	75.2	68.4
$Ave(Rel(X,A))$	0.335	0.210
$Ave(Rel(X,B))$	0.068	0.035
$Ave(Rel(X,C))$	0.002	0.001

表2から1780組の評価データに比べて、動詞評価用の精度が悪いことがわかる。関連が深いと考えられる概念Xと概念Aの平均関連度が0.1以上低くなっている。また、関連があると考えられる概念Xと概念Bの平均関連度においても約半分になっている。ここからも動詞概念における関連度が適切に取れていないことがわかる。

3. Webから自動構築した大規模格フレームを用いた動詞概念の追加と精練

3.1. Webから自動構築した大規模格フレーム

Webテキストから自動構築した大規模格フレーム^{[9],[10]}とは、用言とそれに関係する名詞を用言の各用法ごとに整理したものである。この格フレームは、Web上の約5億文の日本語テキストから自動的に構築しており、約5万用言からなるものである。

3.2. 重み付け手法

ここでは、新たに追加された動詞の属性に対する重み付け手法として、 $if \cdot idf$ の考え方に基づく手法について述べる。

$$W = if \times idf$$

この考え方を概念ベースで利用するため、 if 、 idf を次のように定義する。まず、 if には関連度を用いる。次に、 idf は概念ベースを仮想記事群として新たな idf を算出し、これを以下の式で定義する。

$$idf_{cb(n)}(a_i) = \log \frac{V_{ALL}}{V(a_i)}$$

V_{ALL} は概念ベースに定義される全概念数、 $V(a_i)$ は概念 a_i をn次属性内に持つ概念数である。この式により求められる $idf_{cb(n)}(a_i)$ を概念 a_i のn次概念ベース idf と定義

する。本稿では、3次概念ベースidfの平方根をidfとして用いたとき、最も適切に重み付けができるとされているため、概念 α の属性に概念 β を追加したときの重み $W(\alpha, \beta)$ を、次の式を用いて求める。Rel(α, β)は概念 α と概念 β との関連度である。

$$W(\alpha, \beta) = \text{Rel}(\alpha, \beta) \times \sqrt{\text{idf}_{cb(3)}(\beta)}$$

3.3. 動詞概念の追加と精練

動詞概念は名詞概念と比較して精練されていないため、格フレームを用いて、動詞概念の属性を拡張する手法を提案する。格フレームは、Web上で一般的に語が多く存在し、動詞の属性として適切な語を取得できると考えた。動詞を用いられる格フレームの用言に対応する概念に名詞を追加した。属性の追加語数は最大500とした。しかし、取得される属性の追加候補は、約2500万組と膨大な量があり、雑音も含まれるため、以下3つの方法を用いて選別を行う。それらの属性を精練することによってより適切な連想が可能となる。

1. 従来の概念ベースを用いた関連度を用いて、動詞概念と拡張した属性との関連度を求める。関連度に閾値(0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5)を設け、閾値以上の候補を追加する。
2. 概念ベース上での情報価値も考慮するため、新属性に付与する属性重みに閾値(0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5)を設け、閾値以上の候補を追加する。
3. 動詞の属性の質を向上させるため、格フレームの情報を利用し、格フレームでの検索時に取得される頻度に閾値(1, 10, 100, 1000, 10000, 100000)を設け、閾値以上の候補を追加する。

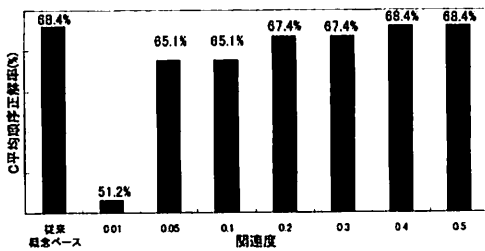


図2. 関連度による閾値実験

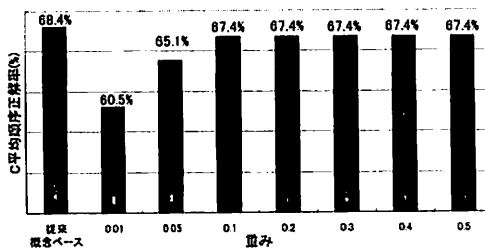


図3. 重みによる閾値実験

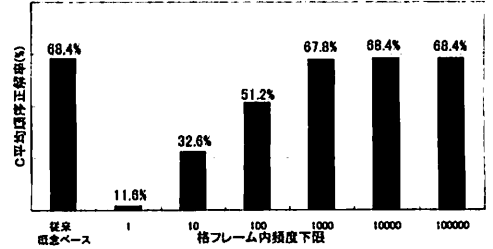


図4. 格フレーム内頻度による閾値実験

図2, 図3, 図4より、いずれの手法もC平均順序正解率は向上せず、逆に低下した。追加する属性が増加したとき、C平均順序正解率が低下した原因として、共通属性の増加が挙げられる。本章では、名詞概念と比較して適切な属性が付与されていない動詞概念の精練を目的に、格フレームを用いて、用言から検索される名詞概念を動詞概念に属性の追加を行った。しかし、この手法では概念ベースの質が向上しないことがわかった。また、その原因がどの用言からも検索できる概念が属性として追加され、不適切な共通属性が増加してしまう点と、動詞概念に適切な属性が付与されていないため、動詞概念を特徴付ける概念であっても高い関連度が算出されない点の2点にあると考えた。これらの解決策として、概念ベースから属性を取得することを考えた。これにより、前者のように不適切な雑音が多くなると考えられる。また、概念ベースから属性を取得するために関連度計算はある程度可能になるだろうと考えられる。次章では、概念ベースからの属性の追加手法を提案する。

4. 概念ベースを用いた動詞概念の追加と精練

4.1. 動詞概念の追加と精練

我々が普段利用する動詞は殆ど漢字で表されることが多く、その漢字によって動詞を特徴づけている。例えば、「走る」という動詞について考えてみる。「走る」は、「走」という漢字によって特徴づけられていると考えられる。動詞の漢字表記されている熟語を概念ベースから検索し、追加をしていった。また、属性を追加する際に概念ベースidf値を用いた。概念ベースidf値とは、3.2節で述べたように、概念ベース内を一つの文書空間とみなすことにより、概念ベース内における頻度を値として取得できるものである。値の低い概念、つまり出現頻度の高いものを追加した。3節より、むやみに属性を取得するだけでは、精度が逆に落ちてしまうことがわかっている。そこで、今回は、1語の動詞につきidf値の低いものの上位15語の概念ベース内に定義されている熟語を追加した。これだけでは動詞を特徴づける属性として少ないということ、また、関連度計算では共通属性が優先的に計算されると

いう、2点が考えられるため、属性として追加した概念の属性も追加した。これにより、動詞概念を特徴づけられると考えた。概念ベースから追加した属性であるので、関連度計算が可能である。しかし、従来から動詞との関連度が低いという問題点があるので、動詞の名詞化を行った。例えば、「走る」という概念を名詞化すると、「走り」というように語尾を変換する。また、「走行する」のようなサ変名詞については「する」を取り除き、「走行」というようにした。

表3. 動詞の名詞化と関連度例

	走る				
属性	走行	競走	馳駆	駆動	走者
関連度	0.536	0.096	0.206	0.004	0.054

表3より、競走、走者、のような、属性として正しい概念についても関連度が低くなっていることがわかる。また、前に示した名詞化手法では、「見る」や「放す」のような語に対して名詞化を行うことが難しいということがわかった。そこで、概念ベースから追加した属性であるので、属性同士で関連度計算を行うことができ、ある程度の関連度が取得できると考えた。そこで、ここでは属性同士で関連度計算を行い、相互関連度を取得した。その相互関連度を利用することにより、追加した属性において無関係である属性を選別できると考えた。相互関連度例を表4に示す。

表4. 相互関連度例

	走る				
属性	競走	御馳走	走者	走行	独走
関連度(競走)	-	0.006	0.117	0.235	0.246
関連度(御馳走)		-	0.005	0.006	0.014
関連度(走者)			-	0.054	0.057
:	:	:	:	:	:

そこで、この相互関連度と重みを用いて以下のような閾値を設けた。

1. 相互関連度を取得し、閾値(0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5)を設け、閾値以上の関連度を5語以上取得できた候補を追加した。
2. 概念ベース内の情報価値も考慮するため、重みに閾値(0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5)を備え、閾値以上の候補を追加する

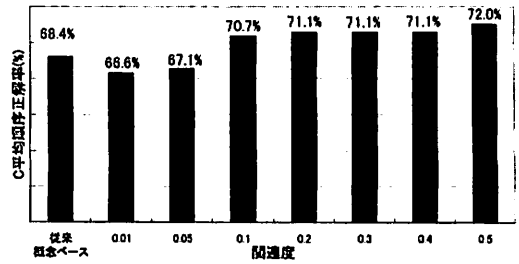


図5. 相互関連度による閾値実験

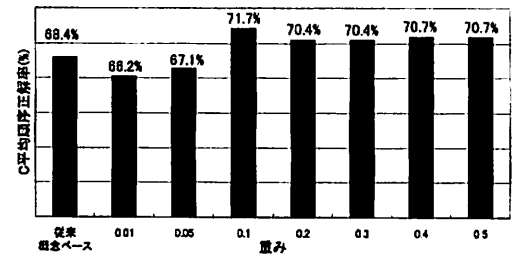


図6. 重みによる閾値実験

図5, 図6より関連度による精練において、閾値を0.1以上に設定することによってC平均順序正解率が向上したことがわかる。また、重みについても閾値を0.1以上に設定することによってC平均順序正解率が向上した。特に、1780組の評価セットにおいても、関連度の閾値を0.1に設定することで0.2%の精度向上がみられた。関連度、重みによる閾値が0.1のときには、XA, XBの共通属性数が増えているのに対し、XCにおいては共通属性数が減っていた。これにより、XA, XBの関連度の値が上がり、XCの関連度の値が下がる、つまり、C平均順序正解率が向上したのだと考えられる。の結果を考慮して、重みの閾値を0.1、関連度の閾値を0.1として2つの閾値を設け新たな概念ベースを作成し、評価を取った。評価結果を表5に示す。

表5. 新概念ベース評価結果

	1780組	動詞評価用
C平均順序正解率(%)	75.4	71.9
Ave(Rel(X,A))	0.340	0.302
Ave(Rel(X,B))	0.070	0.050
Ave(Rel(X,C))	0.002	0.002

表5より、表2と比較すると、1780組評価セットにおいて0.2%の精度の向上が見られた。また、動詞評価用セットにおいては3.5%の精度向上に成功した。今回の動詞概念の属性の追加、精練を行うことにより関連度の値が高く算出されるようになった。XABC評価において、従来の概念ベースでは失敗したが、新概念

ベースでは成功した例についてみていく。

表 6. 新概念ベース成功例

X	A	B	C
	眠る	ベッド	野球
寝る (従来概念ベース)	0.196	0.197	0.001
寝る (新概念ベース)	0.343	0.261	0.001
	記す	控える	吹き飛ばす
書く (従来概念ベース)	0.053	0.076	0.002
書く (新概念ベース)	0.192	0.132	0.004

表 6 より、「眠る」のように我々人間が関連が深いと考えられる語において従来の概念ベースでは、関連度の値が低く算出されている。しかし、今回の動詞概念の属性の追加、精練を行うことにより関連度の値が高く算出されるようになった。「寝る」という漢字表記からだけでなく、その属性を取得することによりこのような結果が得ることが出来たのだと考えられる。また、従来の概念ベースでは正解であった組に対して、新概念ベースでは不正解と判断された失敗例について表 7 に示す。

表 7. 新概念ベース失敗例

X	A	B	C
	音	スタジオ	母ね当てる
音響 (従来概念ベース)	0.284	0.002	0.001
音響 (新概念ベース)	0.921	0.002	0.002
	占い	土地	摩り替わる
風水 (従来概念ベース)	0.022	0.09	0.001
風水 (新概念ベース)	0.022	0.032	0.001

表 7 より、(X - A,B,C) = (音響-, 音, スタジオ, 母ね当てる) のように XA の関連度の値が上がり、良くなったと考えられるものでも、 $Ave(\text{Rel}(X,C))$ の値が上がってしまったことによる失敗が 7 件あった。従来概念ベースにおける、 $Ave(\text{Rel}(X,C))$ は、0.0015 であったのに対して、新概念ベースでは、 $Ave(\text{Rel}(X,C))$ は、0.0022 となった。そのため従来の概念ベースでは成功したが、新概念ベースでは失敗したと考えられる。また、(X - A,B,C) = (風水 - 占い, 土地, 摩り替わる) のように、XB の関連度の値が上がることによる失敗も 10 件あった。本来なら、XA でも関連度の値が上がるように作成した概念ベースであったが、「風水」と「占い」では漢字表記からでは特徴付けられないような語においてこのような失敗がおこっている。このような語については、他の解決策を検討する必要がある。

5. おわりに

本稿では、動詞からの連想を正しく行うために既存の概念ベースの動詞概念の属性を追加する方式を提案した。

動詞概念の属性追加については、大規模格フレームを用いる方式と概念ベースから追加する方式を提案した。大規模格フレームにおいては、属性数が増えることにより、共通属性が大幅に増加し、関連度が正しく取得できないということがわかった。そこで、本稿では、概念ベースから属性を追加精練するほうが良いということがわかった。また、属性をそのまま追加するのではなく、相互関連度や重みを用いて選別することにより、関連度の値を大きくすることが出来たので、従来の概念ベースよりも動詞の連想が、適切に行えるといえる。しかしながら、新たな問題点も見えてきたので、それに対する解決策も検討する必要があるだろう。これらのことを考慮することによって、よりよい連想システムが実現できるだろうと考えられる。

謝 辞

本研究は文部科学省からの補助を受けた同志社大学の学術フロンティア研究プロジェクトにおける研究の一環として行ったものである。

文 献

- [1] 土屋誠司, 小島一秀, 渡部広一, 河岡司. 常識的判断システムにおける未知語処理方式, 人工知能学会論文誌, 17-6, pp. 667-675(2002).
- [2] 笠原要, 松澤和光, 石川勉. 国語辞書を利用した日常語の類似性判別, 情報処理学会論文誌, 38-7, pp. 1272-1283(1997).
- [3] 渡部広一, 河岡司. 常識的判断のための概念間の関連度評価モデル, 自然言語処理, 8-2, pp. 39-54(2001).
- [4] 井筒大志, 東村貴裕, 渡部広一, 河岡司. 概念ベースを用いた関連度計算方式の精度評価, 信学技報, NLC2001-94, pp. 117-122(2002).
- [5] 池原悟, 宮崎正弘, 白井諭, 横尾昭男, 中岩浩巳, 小倉健太郎, 大山芳史, 林良彦(編). 日本語語彙体系, 岩波書店(1997).
- [6] 奥村紀之, 小島一秀, 渡部広一, 河岡司. 電子化新聞を用いた概念ベースの拡張と属性重み付与方式, 情報処理学会研究報告, 2005-nl-166, pp. 55-62(2005).
- [7] 小島一秀, 渡部広一, 河岡司. 連想システムのための概念ベース構成法-属性信頼度の考え方に基づく属性重みの決定, 自然言語処理, 9-5, pp. 93-110(2002).
- [8] 渡部広一, 奥村紀之, 河岡司. 概念の意味属性と共起情報を用いた関連度計算方式, 自然言語処理, Vol. 13, No. 1, pp. 53-74, (2006).
- [9] 河原大輔, 黒橋禎夫, 高性能計算環境を用いた Web からの大規模格フレーム構築, 情報処理学会 自然言語処理研究会 171-12, pp. 67-73, (2006).
- [10] 河原大輔, 黒橋禎夫, 格フレーム辞書の漸次的自動構築, 自然言語処理, Vol. 12, No. 2, pp. 109-131, (2005).