

# 日常語の意味が判別可能な「概念ベース」技術

デモ9

松澤 和光† 笠原 要† 金杉 友子‡ 永森 千晴‡

† NTT(株)コミュニケーション科学研究所 ‡ NTTアドバンステクノロジー(株)

## 1. はじめに

与えられた知識が不完全でも、常識で補って概略的な答えを導く「アバウト推論」の研究[1]の一環として、基本常識たる「概念ベース」[2]を構築し、研究してきた。この概念ベースを用いた類似性判別技術と、言葉遊びへの応用[3]について説明する。

## 2. 概念ベースと類似性判別技術

概念ベースは、言葉の意味(概念)を知識ベースとして体系化したものである。概念は、その特徴を表す属性と重要度の組の集合で表現される。例えば、概念「馬」=(動物,0.5),(家畜,0.3)...となる。属性は電子化された国語辞書等から完全に機械的に獲得したものを精練化し、現在日常語約4万概念について総200万属性という大規模なものが構築済である。

この概念ベースを用い、各概念間の類似性を判別する機構を開発した。これは属性を2800種のカテゴリに分類した上でこれらベクトルのなす余弦によって概念間の類似度を計算する機構である。(図1)

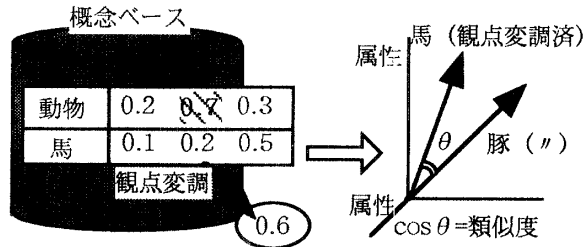


図1：概念ベースを用いた類似性判別

更にこの概念間類似度を、概念が使われている状況に応じて計算する仕組みを開発した。「状況」は概念ベース中の任意の概念あるいはその組み合わせで指定し、これを「観点」と呼ぶ。ある観点Kの下での概念Aと概念Bの類似度は、Kの属性表現中で重要度の大きい属性について、AとBの重要度を各々強調して計算する(観点変調)。これにより図2に示すような観点の違いに応じた類似度を計算することができる。また同様に、ある観点の下である概念に似た概念の検索も行える。

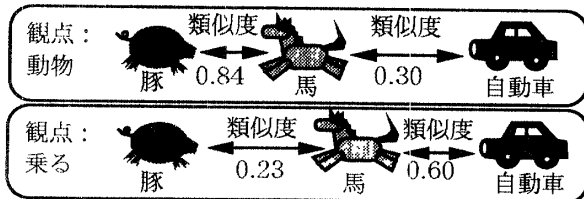


図2：類似性判別結果の一例

## 3. 「言葉遊び」への応用

前述の概念ベース技術により、(1)ある語に表記が似た語を検索する類別機能(表記)、(2)ある語に意味が似た語を検索する類別機能(意味)(3)ある語に関連する語を検索する連想機能が得られる。これらの機能を用い、人間の行う言葉遊びの自動回答を行った。

・クロスワードパズル:

ヒントの単語に関連する/意味の似た語を答として検索し回答。

・連想しりとり: 「バナナといたら黄色」

(単語Aといたら単語B)単語Aから連想する語を連想機能によって検索し、単語Bを回答した。

・なぞなぞ: 「くりはくりでもぎよっと驚くりは?」

(単語AはAでも単語集合Cに関連するAは?)

まず、単語Aの表記に似た語を検索し、その中で単語集合Cに関連する語を検索して回答。

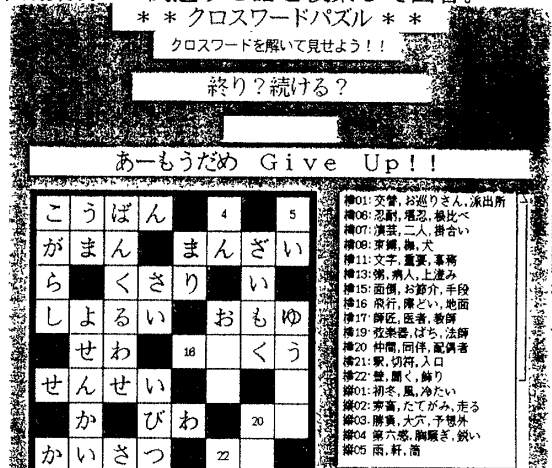


図3：言葉遊びへの応用例

## 4. おわりに

本稿では、概念ベース技術の概要とその技術の応用例を記した。今後も概念ベース技術を利用した様々なシステムを作成し、その有効性の評価を行う予定である。また、当講演で行われるデモンストレーションも併せてご覧いただきたい。

## References

- [1] 松澤和光,湯川高志,笠原要,藤本和則: アバウト推論技術-常識を持ったコンピュータの実現を目指して-, NTT R&D, Vol.45, No.11(1996).
- [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, Vol.38, No.7(1997).
- [3] 金杉友子,松澤和光,笠原要: アバウト推論の「言葉遊び」への適用, 信学技報, NLC96-31, pp.1-7(1996).z