

4C-2 ファジ理論によるあいまいな言葉の処理

小倉久和¹ 高濱徹行¹ 小林 浩²

1 福井大学 2 アトラス情報サービス(株)

1. はじめに

自然言語処理・自然言語理解において自然言語のもつあいまいな性格をどのように扱うかは一つの重要な課題である。このような課題は多くの分野での共通の課題でもある。たとえば、データベースの検索、コンピュータシステムの操作、対話型のエキスパートシステムなどにおいて知的に振舞うマンマシン・インタフェース、あるいは、画像認識におけるパターンや特徴の表現におけるあいまいさの扱い、等々、多くの重要な課題がある。

本研究ではモデル的な対象を仮定し、あいまいな言語表現をファジ理論によって数量化して扱うことを試みた。入力されたあいまいな言語表現をファジ表現に翻訳し、処理し応答する、簡単なシステムを構成した。表現に用いる単語や構文をある程度制限しても余り不自然にならないようにするため、モデル世界は積木の世界とする。位置や大きさ、色などの積木の特徴をあいまいな言葉を含む日本語を用いてシステムに与えると、システムはそれを知識として蓄積する。積木の特徴に関する問い合わせに対して、蓄積した知識を用いて必要な論理演算や推論によって日本語で返答する。

2. モデル世界

本研究の目的はあいまいな言葉をいかにして計算機で扱うかを中心課題としているので、積木の世界を記述する構文と単語を簡単のため大幅に制限する。

単語

- (a) <名詞>: 個々の積木を区別するための呼び名。
 (b) <属性名>と<形容詞>: <属性名>とは“大きさ”や“重さ”などの「積木の特徴」の区分を表わす言葉と、それぞれの<属性名> に対応する<形容詞>。「前後」、「左右」、「大きさ」、「長さ」、「重さ」、「位置」、「色」の各属性と対応する形容詞を扱う。「位置」は「前後」と「左右」の組合せで、「色」はRGBの強度の組合せで扱う。
 (c) <修飾詞>
 タイプ1: “とても”, “大変”, “やや”, “少し”
 タイプ2: “...でない”
 (d) <接続詞>: “さらに”

構文

システムの機能は、モデル世界に関する知識の獲得と、獲得した知識の利用とに分けられる。それに対応してシステムで処理する文章は、システムに対して積木の特徴を与

える教示文と、システムの持つ知識に対して問い合わせをする質問文とする。

(a) 教示文

<名詞>の<属性名>は<比較対象>より<特徴表現>です
 <比較対象>は既知の積木と比較して表現するとき、その比較対象の積木の呼び名を表わす。<特徴表現>は、“とても軽い”のように、<形容詞>に<修飾詞>を伴った言葉である。“大きい”などのように<形容詞>だけでその<属性名>が判断できる場合は<属性名>を省略できる。

<比較対象>を伴わない積木の特徴の表現を「絶対表現」、既知の積木(<比較対象>)と比較して特徴を表わす表現を「相対表現」と呼ぶ。<比較対象>は絶対表現では不要である。

文章中の<名詞>を除く<属性名>、<比較対象>、<特徴表現>より構成される部分を「節」とよぶ。節は“さらに”などの接続詞を伴って1入力文中に複数含むことができる。ただし、<名詞>は1入力文中に複数あつてはならない。1つの教示文ではある1つの積木についての特徴を表わす。複数の特徴を絶対表現、相対表現を組合せて表現できる。

品詞の順序や助詞が異なる表現も、同じ意味を持つならば教示文として可能とする。

(b) 質問文: 次の3つの型を区別する。

① <名詞>の<属性名>は<比較対象>より<特徴表現>ですか?

例: 「積木2の大きさは積木1よりとても大きいですか?」
 <名詞>の積木の特徴とシステムが記憶している特徴とが一致するかどうかを問う質問文である。同じ意味を持つ別の表現も可能としている。

② <比較対象>と比べて<名詞>の<属性名>は?

例: 「積木2と比べて積木1の大きさは?」
 <名詞>で表わされる積木の<属性名>の<特徴表現>を問う質問文で、システムが記憶している積木の特徴を答えるときに用いる。

③ <比較対象>と比べて<属性名>が<特徴表現>の積木は?

例: 「積木2と比べて大きさが大きい積木は?」
 システムが知識として記憶している個々の積木の特徴から、条件に適合する積木を問うときに用いる。

質問文においても<接続詞>を用いて複数の特徴(節)を言い表すことが可能である。

構文規則

すべての構文規則は次の形式をとる。

[非終端記号]=[終端記号]+[非終端記号]

[終端記号]は1つの単語を、[非終端記号]は[終端記号]に続く残りの単語の組みである。文章の初めと終りにもそれぞれ1つの非終端記号を割り当てる。これは文法としては正規文法に属する。

3. システム構成

システム全体は、日本語解析部(入力文の解析)、獲得知識処理部(ファジイ表現への変換)、質問応答部(ファジイ表現から日本語への変換)、およびシステムコントロール部からなる。構文規則、辞書獲得知識、ファジイ表現翻訳知識、日本語翻訳知識は知識ベースとして構成されている。システムコントロール部はこのシステムの核となる部分で、オペレータからの入力と、他の3つの処理部の制御を一括して行なう。他の3つの処理部はシステムコントロール部の子プロセスとして起動されるが、各部は独立したプロセスで、必要な知識ベースに独自にアクセスできる。システムコントロール部と他のプロセスとのコミュニケーションは、共有メモリと次にメッセージ機能を用いて行なう。

4. あいまいな言葉のファジイ表現

日本語解析知識

属性についてのあいまいな言葉、つまり“大きい”などの形容詞は、システムの内部表現として、ファジイ集合で表現する。全てのファジイ表現に共通の表現を与えるため、全ての属性のスケールを0~10の11個の整数値とした。したがって、その属性に対応する形容詞のファジイ表現は、11個の離散的な点に対してメンバシップ関数を与えられたファジイ集合である。

〈修飾詞〉は形容詞で表わされるファジイ集合を変換する演算子とみなす。

タイプ1 (a)集中化 “とても” : 2乗, “大変” : 3乗

(b)拡大化 “やや” : 平方根, “少し” : 3乗根

タイプ2 (c)否定 “...でない” : 補集合

複数の修飾詞がある場合の演算の順序は、表現の意味から考えて、タイプ1は形容詞に近い方から、タイプ2の演算はタイプ1の演算の後演算をする。たとえば、“とても 大変 大きい でない”は“(とても (大変 大きい)) でない”と解釈する。

相対表現において与えられる積木の相対的な特徴は、ファジイ関係およびファジイ関係における演算として表現する。具体的な特徴は、比較対象の積木の特徴とファジイ関係から推論する必要がある。この推論規則は、

事実: x は X である	x, y は積木の名称、 X, Y はその特徴を表わすファジイ集合、 R はファジイ関係
関係: y は x と R である	
結果: y は Y である	

であるから、マックス・ミニ演算の合成規則を用いて

$$Y = R \circ X$$

によって推論する。推論によって得られたファジイ集合が非正規である場合は正規ファジイ集合に変換する。

獲得知識

獲得した各積木の特徴に関する知識は、その各属性を、ファジイ集合、ファジイ関係に対する修飾詞による演算や推論等を用いてファジイ集合として表現され、積木ごとに

フレームの形で各属性をスロットとして蓄積する。知識として与えられていない属性に関しては未定義としておく。獲得知識は、システム動作時には共有メモリに置かれる。

可能性理論による応答

質問応答部では与えられた質問文の特徴が獲得知識における積木の特徴との一致度を求め、それに対応した応答を行なう。2つのファジイ集合の間の一致度は、マックス・ミニ演算によって得られる可能性測度とする。

応答は質問文の型によって次のようになる。

① 質問文の条件の正否を問う

可能性測度が0.5以上であれば“はい”、未満ならば“いいえ”と応答する。複数の節をもつ質問文に対しては、各節の可能性測度の最小値で判定する。

② 積木の〈属性名〉がどんな〈形容詞〉で表現できるか問う

〈属性名〉に対応する〈形容詞〉のファジイ集合との可能性測度の最も大きい〈形容詞〉を採用して、応答する。

③ 獲得知識として持っている積木のうちでどの積木が最も条件に適合するかを問う

獲得知識のそれぞれの積木に対して(1)と同じようにして最大の可能性測度の積木を応答する。

5. 今後の課題

(1) まず第一に、システムで扱える日本語の構文と単語を拡張し、より自然な文章が処理できるように改善する必要がある。本システムでは、正規文法あるいは文脈自由文法を前提としている。この範囲でもっと自然な文章を処理できるようにすることが可能である。

(2) あいまいな感覚は人それぞれに異なっており、同じファジイ集合だけで表わすことには無理がある。オペレータの感覚を理解しそれにシステムを追随させていくようにすることは、適応制御の考え方にも関連すると思われる。

(3) 獲得知識は互に独立な知識として蓄積される。しかし、それらは互に無関係ではないし、大量の知識が蓄積されたときには何らかの組織化が避けられない。

(4) 可能性理論による応答ではマックス・ミニ演算によって可能性測度を求めるため、内容の異なるファジイ集合に対して同じ可能性測度が得られることがある。きめの細かい応答が必要な時には障害になるかもしれない。

(5) 最後の応答の生成は極めて簡単な機構によって形式的な応答しかできていないので、日本語翻訳のための規則を知識ベースとして拡充していく必要がある。

(6) 知的に振舞うシステムとしては、未知の情報に対する推測、推論、解決、あるいは学習や記憶などに対する処理が不可欠であるが、本システムでどのように考えていくか、特にファジイ論理との関連ではまったく未知の世界である。

参考文献

- 馬野元秀：ファジイ集合の概念を用いた日本語質問応答システムの拡張，情報処理学会第37回全国大会予稿集(1988)、
馬野元秀：ファジイ集合論とソフトウェア，情報処理，Vol.30, No.8, p922, (1989)
銀林純，藤井則夫：「三才児」の日本語知的インタフェース，情報処理学会第37回全国大会予稿集(1988)