

6H-8

ファジィ集合の概念を用いた
日本語質問応答システムの拡張

- ファジィ推論の導入 -

馬野 元秀

(大阪大学 大型計算機センター)

1. はじめに

我々は、以前、あいまいな意味を持つ単語を用いることのできる簡単な日本語質問応答システムを試作した[1]。今回、これを拡張して、ファジィ推論を行なう機能を追加したので、報告する。

2. 従来のシステムの概要

従来のシステムは、文献[2]の小型日本語質問応答システム TQAS を VAX-11/780 (OS は VAX/VMS) 上の Lisp-F3 (Interlisp のサブセット) に移植したものに、あいまいな概念を含む単語を処理する部分を追加して実現した。そして、例えば、次のような会話が可能である。

- > shiroi hakowa chiisai.
(白い箱は小さい)
- * wakarimashita.
(わかりました)
- > shiroi hakowa totemo chiisai desuka?
(白い箱はとても小さいですか?)
- * hai tabun totemo chiisai deshou.
(はい、たぶんとても小さいでしょう)
- > aoi hakowa ookisaga chukurai desu.
(青い箱は大きさが中くらいです)
- * wakarimashita.
(わかりました)
- > shiroi hakowa aoi hakoyori zutto chiisai desuka?
(白い箱は青い箱よりずっと小さいですか?)
- * hai tabun zutto chiisai deshou.
(はい、たぶんずっと小さいでしょう)

> で始まる行が入力行で、* で始まる行がシステムからの応答である。現在のところローマ字しか使えないので、本報告では、漢字かな混じり文を補うことにする。

システム内部では、あいまいな概念を可能性分布で表現している。今は、白い箱と青い箱の大きさは、

白: {1/0, 0.92/1, 0.68/2, 0.32/3, 0.08/4}
青: {0.222/3, 0.778/4, 1/5, 0.778/6, 0.222/7}

となる。質問に対しては、質問中のファジィ集合に対する事実の可能性分布の満足度を計算し、これに基づいて返答している[1]。

取り扱える自然言語の範囲は狭く、名詞は、hako(箱)、hana(花)、mari(まり)、tsukue(机)、tsumiki(積木)の5つだけである。属性には、「色」、「大きさ」、「重さ」、「長さ」がある。そして、「色」の属性で使用できる形容詞は、akai(赤い)、aoi(青い)、kiiroi(黄色い)、kuroi(黒い)、shiroi(白い)の5つで、これらはあいまいではないと考えている。また、「大きさ」の属性で使える形容詞は、ookii(大きい)、chukurai(中くらい)、chiisai(小さい)であり、「重さ」では、omoi(重い)、chukurai(中くらい)、karui(軽い)であり、「長さ」では、nagai(長い)、chukurai(中くらい)、mijikai(短い)である。これらは0から10までの11段階の尺度で考えているので、nkurai(n くらい: 1 < n < 9)も使うことができる。さらに、副詞として、色以外の属性の形容詞に対して使える totemo(とても)、taihen(大変)と maa(まあ)、sukoshi(少し)とがある。

3. ファジィ推論の導入

ある箱の大きさを言うのにすでに大きさが分かっている箱と比較して言いたいことがある。例えば、「黒い箱は青い箱よりずっと大きい」などである。

いままでのシステムでも、「白い箱は青い箱よりずっと小さいですか?」のような2つの対象に関する質問は可能であった。これは白い箱の大きさ(これはすでに分かっている)と青い箱の大きさ(これもすでに分かっている)を比べて、どの程度「ずっと小さいか」によって回答することができた。

しかし、「黒い箱は青い箱よりずっと大きい」などの場合には、青い箱の大きさ(例えば、「中くらい」の大きさ)と「ずっと大きい」というファジィ関係から

黒い箱の大きさを推論する必要がある。これは、質問の場合とは処理のしかたがまったく異なり、従来のシステムでは実現されていなかった。

このような推論の方法は、Zadeh により近似推論 (approximate reasoning) として提案されている[3]。いまの場合は、A と B を可能性分布またはファジィ集合とし、R をファジィ関係とするとときに、推論形式が

事実：x は A である
 関係：y は x と R である
 結果：y は B である

であるので、

$$B = R \circ A = \{ \max_{u \in U} (\mu_R(v, u) \wedge \mu_A(u)) / v : v \in V \}$$

となる(推論の合成規則)。

そして、ファジィ関係として、「より大きい」、「より小さい」、「少し大きい」、「少し小さい」、「ずっと大きい」、「ずっと小さい」、「ほぼ等しい」を 使えるようにした(図1を参照)。この結果、次のような会話が可能になった。

> kuroi hakoga arimasu.
 (黒い箱があります)
 * wakarimashita.
 (わかりました)
 > kuroi hakowa aoi hakoyori zutto ookii.
 (黒い箱は青い箱よりずっと大きい)
 * wakarimashita.
 (わかりました)
 > kiroi hakoga arimasu.
 (黄色い箱があります)
 * wakarimashita.
 (わかりました)

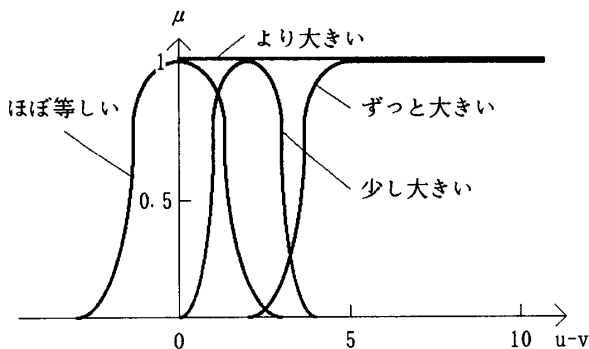


図1. ファジィ関係

> kiroi hakowa shiroi hakoto ookisaga hobo hitoshii.
 (黄色い箱は白い箱と大きさがほぼ等しい)
 * wakarimashita.
 (わかりました)
 > akai hakoga arimasu.
 (赤い箱があります)
 * wakarimashita.
 (わかりました)
 > akai hakowa aoi hakoyori ookii.
 (赤い箱は青い箱より小さい)
 * wakarimashita.
 (わかりました)

なお、現在ところ、この例のようにまず箱が存在することをいっておく必要がある。この結果、それぞれの箱の大きさは

黒：{0.222/6, 0.222/7, 0.778/8, 0.778/9, 1/10}
 黄：{1/0, 0.92/1, 0.778/2, 0.68/3, 0.32/4, 0.222/5, 0.08/6}
 赤：{0.222/4, 0.778/5, 1/6, 1/7, 1/8, 1/9, 1/10}

となった。したがって、

> kuroi hakowa ookii desuka?
 (黒い箱は大きいですか)
 * hai tabun ookii deshou.
 (はい、たぶん大きいでしょう)

などのような質問に答えることができる。

今回の拡張に際して、システムを大阪大学大型計算機センターの ACOS-2000 (OS は ACOS-6/MVX) 上の UTI-Lisp に移植し、ファジィ集合の処理を「ファジィ集合処理システム」を使用するように書きかえた。

4. おわりに

今後の拡張としては、取り扱える日本語の範囲の拡張と推論で得られた結果を適切な言葉で表現し直す言語近似の実現を行ないたい。

[参考文献]

1. 馬野 (1985): 「ファジィ集合の概念を用いた日本語質問応答システムの試作」、情報処理学会研究報告、Vol. 85, No. 36, pp. 81-88, A142-11 (知識工学と人工知能研究会)。
2. 田中、元吉、山梨(1983): 「LISP で学ぶ認知心理学 3 言語理解」、東京大学出版会。
3. L. A. Zadeh (1975): "The Concept of a Linguistic Variable and Its Application to Approximate Reasoning," Information Sciences, Vol. 8, pp. 199-248, Vol. 8, pp. 301-357, Vol. 9, pp. 43-80.