

助数詞の相互学習システム

2C-8

黎天舒 豊田幸雄 山口昌也 乾伸雄 小谷善行

（東京農工大学 工学部 電子情報工学科）

1. はじめに

我々はあるものを数えるとき、通常助数詞を使う。例えば、魚を数えるとき、一匹、二匹といい、牛を数えるとき、一頭、二頭という。日本語では、助数詞は数詞と接続して名詞の数を表すが、どんな名詞にどんな助数詞を使うかに一般的な規則が存在する。システムにその規則を持たせて、ある名詞に対して使われる助数詞を自動的に推論するのが本システムの基本的な考えである。それに基づいて、「助数詞の相互学習システム」を実現した。

本システムは利用者とシステムの両方の知識が成長していくことを目的とする。システムは獲得した知識を利用者に教える。利用者は具体的な言葉に使う助数詞から、ほかの助数詞に使う単語を連想して、自分が持つ助数詞についての知識をどんどん整理して、システムに教える。それによって、利用者の知識が再構成、再確認する[1]と同時にシステムも新しい知識を獲得することができる。

2. 日本語における助数詞

日本語においては、一般的に一つの助数詞が特定の意味属性を持った名詞に対して使われる。例えば、「本」という助数詞が表すものとして、ペン、木、傘、タバコ、ネギなどのような細長いものが挙げられる。もし「細長い」という意味属性を持った名詞に助数詞「本」を使うことを一般的規則と見なせば、「髪の毛」にきつと「本」を使うことができるだろう。このような助数詞と名詞との対応関係の例を表1に示す。[2]

もちろん、人間の言葉は自然発生したものなので大枠の規則がある一方、習慣的な言い方も多い。例
Bi-directional Learning System for Counters
Tianshu LI, Yukio TOYODA, Masaya YAMAGUTI, Nobuo INUI, Yoshiyuki KOTANI
Tokyo University of Agric. and Tech., Dept. of Computer Science

えば、同じものに複数の助数詞を使用することや、特別の場合に特別な助数詞を用いることがある。

表1 助数詞と名詞との対応関係の例

名詞の意味属性		助数詞	実例
細長いもの		本	ペン、木
動物	鳥類	羽	ツバメ
	小さい動物	匹	魚、虫
	大きい動物	頭	牛、馬
薄くて平たいもの		枚	紙、皿

3. システムの設計について

3.1 概念階層の構築

システムは意味属性を概念階層構造の形で持つ。例えば、動物においては、鳥類に使う助数詞は「羽」であって、それ以外では、魚、虫などのような小さい動物に「匹」を、牛のような大きい動物に「頭」をよく使う。したがって、「動物」という基本的な意味属性に対して、もっと細かく分類すれば、概念階層ができる。そのようにすると、複雑な助数詞の個々の小さい範疇が分かる。これらを用いて、助数詞を正確に推論することができる。

3.2 対話による知識の獲得

システムはカウンセリング的な利用者インターフェースを採用して、利用者が具体的な言葉から、多くの言葉を引き出す誘導対話の戦略を取っている。システムは知識ベースを用いて、利用者にある名詞に使う助数詞および名詞の属性情報を提示する。利用者は新たな連想をして、またシステムに教える。このように、利用者はシステムとのやりとりしながら、両方の知識が広がっていく。それは次のような二つの方式を通じて実現した。

(1) システムは話題の主導権を持つ。

システムは知識ベースを用いて利用者に質問する。利用者の答えは知識ベースに保持しているものと一致するとき、ほかの話題を遷移する。一致しないと

きさらに原因を分析する。同じものに二つ以上の助数詞がよく使われる。例えば、「匹」も「尾」も魚に関する助数詞である。システムはこのような場合を扱うとき、推論した助数詞のほかに利用者の答えと一緒に知識ベースに登録する。また、特殊な言語習慣により名詞で助数詞を分類することが困難なとき、助数詞の特例として扱う。

(2) 質問する主導権は利用者にあげる。

利用者の質問に対して、答えが存在しないとき、システムが概念階層の属性情報を提示して、利用者を選ばせる。そしてそれを辿って、助数詞を推論する。もし対応する意味属性が概念階層に含まれていないならば、システムは最上位概念に意味属性を追加する。例えば、システムは「計算機」に使う助数詞は「台」であることを獲得した。もしまた「計算機のほかにどんなものに「台」を使いますか」と反問したら、利用者は更に「機械」という意味属性を教えて、新しい意味属性は階層構造に立てられる。

4. 実現例

実現したシステムの対話例の一部分を載せる。

システム：まず、私の質問を教えてください。

ユーザ：はい。

システム：ツバメに使う助数詞は何ですか。

ユーザ：羽。

システム：はい、そうです。そして「羽」は鳥類を数えるときよく使います。

...

システム：あなたが質問して、私が答えます。

ユーザ：はい。

システム：何の助数詞が聞きたいですか。

ユーザ：豚。

システム：私は答えられません。

この言葉は次の意味属性を持っていますか。(あったら入力してください)

1. 動物
2. 細長いもの
3. 薄くて平たいもの

ユーザ：動物。

システム：それに対する助数詞は羽ですか。

ユーザ：いいえ。

システム：それに対する助数詞は匹ですか。

ユーザ：いいえ。

システム：それに対する助数詞は頭ですか。

ユーザ：はい。

システム：豚を数える時、一頭、二頭...と言いますね。分かりました。

以上の対話のように、システムは既存のデータから出題して(ツバメ)、利用者の答え(羽)と一致すると、それについての情報を説明する。利用者に対して、勉強になったと同時に、さらに思考が誘導される。続いて、主導権を利用者に渡って、利用者が頭に思い浮かべたもの(豚)から、話題を展開していく。システムは知識を持っていないとき、概念階層の上位から利用者に質問する。その言葉の意味属性(動物)が分かったあと、それに関する助数詞を推論する。システムが提供した個々の仮説を利用者は選択して、最終的に、正確な助数詞を選び出すことができる。これにより新しい知識が獲得される。

5. おわりに

本稿では、利用者がコンピュータに教えるという考えに基づく助数詞相互学習システムについて述べた。助数詞の使用上の規則をつかむことは、日本語を勉強している人に必要なことである。助数詞は日本語だけではなく、中国語にも現れる。言葉使用及び適用範囲は違うところがあるが、共通点もある。今後は今のシステムをもとに、中国語に対応する部分も加えて、中国語と日本語における助数詞の相互学習システムを考えている。

参考文献

- [1]小谷善行：IAC-「利用者が教える」というパラダイムによる教育ツール, 情報処理学会「教育におけるコンピュータ利用の新しい方法」シンポジウム論文集, pp. 49-53, 1989.
- [2]三浦昭・マクグロイン花岡直美：語彙, 荒竹出版, 1988.