6E-6

英語数量表現の解析

飯野　香　　亀井　真一郎

日本電気（株）

1. はじめに

数量表現とは、文字どおり数や量を表す表現である。「1」「one」等の数詞や「any」「いくつか」等の語とそれに対する様々な修飾語句を使って言い表す。

このような表現は多様であり、かつ個別的である。翻訳を行う場合に、その多様な表現をそのまま1対1に写像すればよいという考え方もあるが、それは現実的ではないし、また機械翻訳システムPIVOTで目指している多言語間翻訳を考えると不可能であろう。したがって、あらかじめいくつかの基本形を設定しておき、それを相手言語に伝達すると言う形式をとる必要がある。そこで、どのように基本形を設定するか、多様な表現をいかに基本形に対応づけるか、が問題となる。

本稿では、英語数量表現、特に数詞を使った表現について、解析上の問題点を洗い出し、その構文上の特徴から数量表現を解析する方法について検討する。さらに、ここで提案する方針に基づいて、PIVOT上で解析実験を行った結果について報告する。

2. 数量表現の特徴

数量表現を解析する上での問題点として次のようなことがある。

(1) 表現方法が多様である

数量表現はその表現方法が多様であり、関与する言語現象が多い。

例 (1) The shaft length is <u>more than or equal to 5 inches</u>.
(2) I have <u>two or more apples</u>.
(3) I have <u>two apples or more</u>.
(4) <u>The fourth last car</u> is black.
(5) <u>The last car but three</u> is black.

(2) 特殊な意味になる語句がある

数量表現を修飾する、あるいは数量表現に修飾される場合に独特な意味になる語句がある。

例 (6) The shaft is 20-inch <u>long</u>.
（長さ）
(7) He <u>weighs</u> 50 kg.
（重さがある）
(8) The sensitivity of control is on the <u>order</u> of 0.1° C.
（台）

これに対して、解析の手がかりになるような特徴としては次の点があげられる。

(1) 局所的に認定することができる。

数量表現の統語構造は、文の全体を見渡さなくても、数量の直前・直後のある範囲を見て、意味をとることができる場合が多い。

(2) 語彙的に認定できる。

数量表現を形成する語、あるいは数量表現に修飾されて特別な意味を持つ語は、ある程度限られた語彙に属している場合が多い。

(3) 基本パターンの変形として整理できる。

ひとつの数量を表すために多様な表現が存在する。しかし、それらが表現しようとしている対象は数種に分類することができる。また、数量が他の語を修飾する場合も、基本パターンの変形として整理することが可能である。

3. 解析方針

2.で述べたような数量表現の特徴から、数量表現を認定し意味を抽出することは、解析の早い段階でも可能であると考えられる。そして、あらかじめ数量を解析しておくことで、数量表現との係り受け関係に特徴のあるような語の語義の選択に「数量がある」という情報を役立てることができる。

そこでPIVOTでは、構文解析の前の段階、構文解析前処理部において数量表現の意味を抽出しておくことにした。（図1参照）

Analysis of English Numerical Expression
Kaori IINO, Shin'ichirou KAMEI
NEC Corporation

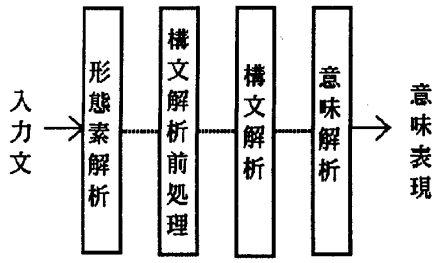

図１　PIVOT英語解析部構成図

このような方針で語彙的に解析を行うためには、数量表現の一部として、あるいは数量表現に修飾されて文中に出現する語の辞書に情報を付与しておくことが必要である。

今回設定した辞書情報の例を図２に示す。

	辞書情報	語例
(1)	数量を修飾して領域を表す	up to from～to～ above under
	数量を修飾して概略を表す	about approximately
(2)	数量に修飾されて属性を表す	long length
	数量に関する操作・状態を表す	average weigh

図２　辞書情報

ここで、(1)は数量の構造・意味を解析するための情報、(2)は数量と他の語の間の意味関係を解析するための情報である。特に、(2)に属するような語は通常使われる意味とは違った意味を持つ場合が多いので、適切に訳し分けるためにもこの情報は重要である。

4.　評価

3で述べたような方針でPIVOT上に数量表現解析規則を作成し解析実験を行った。

出力文の評価結果を図３に示す。

Ａ、Ｂは例文の種類である。Ａは主に数量を表す部分の解析能力を見るための単文中心の例文であり、Ｂは数量表現と他の語句との関係の解析能力を見ることを目的とした例文で複文も含んでいる。それぞれ250文である。

数量表現が正しく解析され、その係り先の語の用法選択（構文解析）・意味選択（意味解析）、それらの間の意味関係が全て正しい場合を解析成功として、解析成功率を算出した。

	A	B
構文解析	１００％	８５％
意味解析	９５％	８０％

図３　解析成功率

図３からわかるように、数量そのものを表す部分はほとんど問題なく解析できることが確認できた。しかし、他の語句との関係の解析はやや解析成功率が落ちる。これは、前置詞や副詞、さらに動詞の格フレーム等の曖昧性のためだと考えられる。

4.　おわりに

数量表現の統語構造の調査・分類を行い、文を意味表現で表すために数量表現中に現れる語句の辞書にどのような情報をもたせるか、どのような方針で解析規則を記述するかについて検討した。また、実験によりこの方針の有効性を確認した。

しかし、表現が複雑になってくると他の語句との関係の解析で誤ることがある。これを解消するには文脈等もっと高次の情報を利用しなければならない部分ももちろんあるが、「数量表現をともなっている」という情報をもっと積極的に利用することによっても、解析成功率の向上を図ることができると考えている。

さらに、今後は数量表現のように語の並びを見て解析を行うような言語現象をどのようにして効率的に扱うかについて、規則の適用方法・システム構成を含めて検討していきたいと考えている。

参考文献

富井篤、他：”英語の決め手数量表現”
インタープレス(1986)
池内正幸：”名詞句の限定表現”
新英文法選書、大修館書店(1985)
亀井、村木：”程度表現のモデル化”
信学 NLC88-6