

# 曖昧な数量詞を含む名詞句の解析法

横山 晶一 落合 武  
山形大学 工学部

## 1 はじめに

日本語の数量詞は、文の中の種々の場所に現れて特殊な振舞いをする、いわゆる浮遊数量詞という現象が知られているが、文中の同じ位置でも曖昧な場合がある。たとえば、「キロ」という数量詞は、重さを示すキログラムと、長さを示すキロメートルのいずれの場合にも用いられ、しかも後者は、長さ、距離、速度のどれを表すのかが曖昧になるケースがある。本論文では、このような曖昧な数量詞をいくつか取り上げ、機械処理のために、数量詞を含む名詞句を解析した結果について述べる。数量詞と共に起る名詞を意味的に分類することにより、これらの数量詞の曖昧性を解消するためのアルゴリズムが作れることを示す [6]。その際、意味分類だけでなく、デフォルト値の設定が有効なことも示す。

## 2 数量詞の分類

### 2.1 数量詞の定義

数量詞は、大きく分けて、単位をしめすものと、数を数えるための補助になるもの（いわゆる助数詞）とがある。ここで取り上げるのは、主として前者に属するものである。どちらも、文の中の任意の位置に現れる、いわゆる浮遊数量詞になることができる（例：「1kgのりんごを買う」、「りんご1kgを買う」、「りんごを1kg買う」）。しかしながら前者は、その単位ではかれるものであれば広い名詞に対して適用できるのに対し、後者は一部の例外（「人」、「個」、「匹」など）を除いては、付随する名詞に限られるものが多い（例：「一冊の本」）。

本稿では、特に前者に属する数量詞を分類して、名詞句の中にその数量詞が出現したときの機械处理的な取り扱いについて検討した。数量詞は上記の浮遊といった現象の他にも、各々の数量詞の現れ方、用いられ方に非常にアドホックな面があるが、ここでは極力その性質を排して、なるべく一般的に取り扱うことができるようなアルゴリズムを作成した。

数量詞の言語学的な取り扱いについては、「つ」と助数詞を比較したもの [3]、浮遊数量詞を扱ったもの [2]、数量詞の意味変化について研究したもの [7] などがある

が、機械処理の観点から数量詞を扱ったものは意外に少なく、わずかに浮遊数量詞を扱った例がある [1] 程度である。

### 2.2 意味による数量詞の分類

数量詞を、意味的に曖昧でない単位、複数の意味を持つ数量詞、助数詞としてもものを数えるときに使用するものの 3 種類に分類した。

#### 1. 意味的に曖昧でない単位：

例：メートル、グラム、ドル、パーセント

#### 2. 複数の意味を持つ数量詞：

cc, リットル (体積) 100cc の水  
(容積) 100cc の瓶  
(排気量) 2000cc の自動車  
度 (角度) 10 度の傾斜  
(アルコール度数) 40 度の酒  
(温度) 40 度のお湯  
(音階) 4 度の関係  
(回数) 3 度の失敗  
インチ (長さ) 10 インチ  
(大きさ) 17 インチのテレビ  
キロ (距離) 自宅から 200 キロ  
(速度) 80 キロで走る車  
(重さ) 10 キロの米  
(容量) 2.5 キロの洗濯機  
ポイント (文字の大きさ)  
10 ポイントの字体  
(得点) 10 ポイントのリード  
(パーセンテージ)  
10 ポイントの増加

#### 3. 助数詞：

例：番、人、個、棹、軒、匹、冊、本、ダース

このうち、本稿で取り扱うのは主として 2 番目のカテゴリに属するものである。以下ではこの種の数量詞の機械的な取り扱いについて考察する。

## 2.3 曖昧な数量詞の現れ方

2番目のカテゴリに含まれる数量詞を用いると、浮遊数量詞ではなくても、次のような曖昧な文が存在する。

- (1) ごみは河口から 10キロにわたって続いていた。
- (2) 雪のため、9日夜から 50キロに規制していた。
- (3) たまご酒を作るとき、まず酒を 50度に温める。
- (4) 水で割って、ウォッカを 10度にした。

(1), (2) の文は、ともに、「～から…キロ」という形を取る。しかしながら、(1) では、キロは距離を示し、(2) では、一般に速度を示す（この文では、きわめて稀な使い方であるが、重さを示すこともある）。(3), (4) も、ともに、形は「～を…度」であるが、(3) では温度、(4) ではアルコール度数を示す。これらをすべて統一的に取り扱うことは不可能である。以下では、数量詞を含む名詞句に着目して、その句パターンを用いて解析する。

## 3 句パターンによる数量詞の曖昧性解消の方法

### 3.1 手法の概要

数量詞を含む名詞句を、まず数量詞、名詞、助詞の位置関係によって、「名詞＋数量詞」（例：幅 5 キロ）、「名詞＋助詞＋数量詞」（例：東京から 50 キロ）、「数量詞＋の＋名詞」（例：30 キロの車）というパターンに分類する。次に、各々のパターンについて、例文を調べた上で、どの数量詞に同定するかの判断基準を、パターン中の名詞の意味を用いて定める。ここで用いる意味は、分類語彙表 [5] と、日本語語彙大系 [4] に示された分類番号を用いる。以下では前者を「分」、後者を「大」で示す。

### 3.2 名詞＋数量詞

数量詞を含む名詞句が現れたときには、まず、「名詞＋数量詞」の形かどうかを判定する。すなわち、次のような場合である。

- (5) 時速 40 キロで走る車が壁に衝突した。

この例文では、数量詞の前に現れる名詞の形だけで、「キロ」＝速度と容易に判定できる。しかしながら、ここでは文のパターンのみを考慮して、文法的なことは考慮に入れていないので、「約 50 キロ」のような名詞句については、名詞の形のみからは判定できない。したがって、この場合には判定を留保し、動詞の形や動詞と共に現れる名詞の形で判定を行う（後述）。ここでは、数量詞の意味を一意に確定する名詞のみを集めて、その分類番

号を判断材料として用いる。これを表の形でまとめたものが表 1 である。

表 1 名詞＋数量詞のときの単位と意味

単位	意味	分類番号	例
キロ	速度	分 1.1940	時速
	長さ	分 1.1920-1, 1.1920-4	幅
	距離	分 1.1731, 1.5260	北
	重さ	分 1.1930-4, 1.1930-5, 1.1900-2	重さ
度	角度	分 1.1930-2, 1.1930-3	北緯
	温度	分 1.1912-10, 1.1960-9	マイナス、華氏

分類番号は、この表ではすべて分類語彙表に記載の番号（この表ではすべて体の類）を用いている。1.1940 が速度、1.1920 が長短・広狭で、その下の 1 は長短、長さなど、4 は内のり、幅、川幅など、1.1731 は方面・方角、1.1930 は角度・軽重・寒暖など、1.1900 は量、1.1912 は整数・対数など、1.1960 は単位・助数接辞をそれぞれ示す。

この表に基づいて、「度」を処理する流れ図を図 1 に示す。図で分かるように、「名詞＋数量詞」という形のときに、名詞の意味を分類番号で判断して、角度か温度かを同定する。キロについても同様の処理を行う。

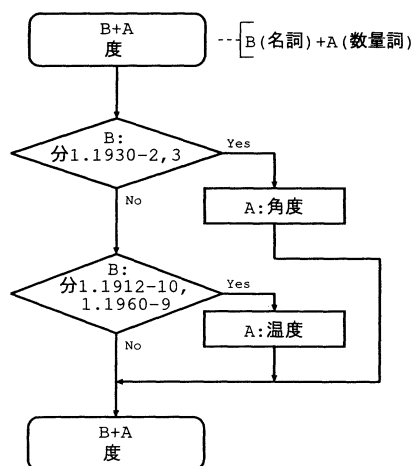


図 1: 「度」処理の流れ図

表1に示したキロ、度以外の数量詞では、cc、リットルは体積、インチについては長さ、ポイントについては得点の意味にはほぼ定まるのでこれをデフォルトにする。

### 3.3 名詞＋助詞＋数量詞

このケースは、名詞と数量詞の間に助詞が入っている場合で、次のように、数量詞「キロ」の場合のみ扱うことができる。

(6) 山形は東京から 350 キロのところにある。

助詞「より」、「から」などを介したキロは、距離を示す。その他の場合には、すでに述べたように、文法的な関係を考慮せずにパターンマッチングのみで判定しているため、必ずしも意味を一意に決定できない。たとえば、(6)と同じパターンを取る構文で、(2)と同様の次の文では、必ずしもキロは距離にならない。

(7) この道路は夜から 50 キロに規制される。

また、B が数量詞となる次のような場合も除外される。

(8) 彼は体重が 55 キロから 65 キロに増えた。

結局、「から」の前が場所の場合には距離、時間の場合にはそれ以外ということになる。

### 3.4 数量詞＋「の」＋名詞

これは、いわゆる「A の B」の形 [8] で、しかも A が数量詞になる場合である。この場合には、それぞれの数量詞に関する例文を調べることによって、表2のように分類できる。

表2の分類番号の部分で、大の次の番号は日本語語彙大系の分類番号を示し、番号の次の語が、その分類の内容である。分の次は分類語彙表の番号で、分 1.1610 は時間、1.1920 は長短・広狭、1.3371 は旅、行楽を示す。また、角度のところの 1.1750 は面・側・表裏、1.1930 は角度・軽重・寒暖などを示す。「度」では、通常は温度と解釈するのが妥当なので、温度をデフォルトとする。

一般的には、表2のような設定で数量詞の曖昧性を解消できるが、特に「キロ」、「度」の場合には、名詞辞書の側に「デフォルト値」といったものを書いておく必要がある。

(9) 80 キロの電車では2時間かかる。

例文(9)では、人間には、「80 キロ」が速度を示すことが容易に分かるが、機械処理のためには、電車の重さと、電車の速度といったものがある範囲を定めて「デフォルト値」として辞書に書いておく必要がある。電車の場合には、重さと速度が数値的に重なることはないが、名詞によっては重なるデフォルト値を書かざるを得ない場合

もある。そのときには複数の解釈を許して、他の条件から判定することになる。

表2 A(数量詞)のB(名詞)場合の単位と意味

単位	意味	分類番号	例
cc, リッ トル	排気量	大 987 乗り物 (本体)	自動車
		大 964 原動機, 「排気量」	エンジン
	容積	大 904 容器, 「容積」	びん
	体積	その他	水
キロ	速度	大 987 乗り物 (本体), 4 人, 536 動物	電車
	長さ	大 417 交通路, 「長さ」	橋
	容積	大 904 容器, 「容積」	漬物桶
	距離	分 1.1610-1, 1.1920,1.3371	旅行
	重さ	大 2 具体, 「重さ」	体重
度	アルコール度数	大 861 酒	ウォッカ
	角度	分 1.1750-8, 1.1930-2	スロープ
	温度	大 1 名詞 [default]	氷
	回数	大 1000 抽象名詞	
ポイ ント	文字の 大きさ	大 1092 文字, 1094 字体, 1095 書体	明朝体
	パーセン テージ	大 2264 増減	増加
	得点	その他	リード

## 4 共起する動詞による判定

上の基準に照らして判定できない場合には、共起する動詞と判定を行う。動詞の分類番号を分類語彙表に基づいて選択した結果、次の表のような結果を得た。分類番号先頭の数字の2は「用の類」、3は「相の類」である。

表3 数量詞と動詞との関係

数量詞	意味	分類番号	動詞の例
cc, リットル	体積		
度	角度	2.1514-1	傾ける
		2.1700-14	前傾する
	温度	3.5150-1～6	熱い
		2.5170-1A,2A	暖まる
ポイント	パーセン テージ	2.1540-2	上昇する
		2.1580-2B	増える
	得点	2.3700-5A,5B	(点を)取る
キロ	速度, 距離	2.1510	運転する
		2.1522	走る
		2.3392-1～5	駆ける
	容積	2.1532-1	入る
	重さ	2.3393-0	持つ
		2.1540-3	持ち上げる

表では、「cc,リットル」のところには、何も示していないが、これはほとんどの場合、動詞に関係なく体積を示すからである。容積や排気量は、むしろ名詞との関連で現れる。「度」は、2.1514（傾斜・転倒など）に属する一部の動詞と、2.1700（位置・方向）に属する一部の動詞が角度を表し、3.5150（気象）と2.5170（熱）に属する動詞、形容詞が温度に関わる。「ポイント」は、2.1540（上がり下がり）と2.1580（増減）がいずれもパーセンテージに関係し、2.3700（所有・取得）に属する動詞が得点に関係する。さらに、「キロ」は、2.1510（動き）、2.1522（走り・飛び・流れなど）の動詞と、2.3392（足の動作）の一部の動詞が速度、距離を表し、2.1532（入る、いれる）に関係する動詞は容積に関わる。また、2.3393（手の動作）と、2.1540（上がり下がり）の一部動詞との共起は、重さを表すことを示す。

一般的には、動詞との共起のみで数量詞の曖昧性が解消できるケースはそれほど多くはない。

## 5 おわりに

なるべくアドホックな取り扱いにならないように、数量詞の曖昧性を解消するアルゴリズムを提案した。しかしながら、文中への数量詞の現れ方はきわめて多様であり、文法的な情報をそれほど用いていないこともあって、上のような解決法では曖昧さの残る場合も多い。

浮遊数量詞については、「数量詞＋の＋名詞＋格助詞」（例：5人の犯人が…）、「名詞＋格助詞＋数量詞」（例：学生を10人…）、「名詞＋数量詞＋格助詞」（例：子供7人と…）といったパターンに対して例文を作成し、成立するか不成立かを調査している。その際に、数量詞と名詞とが同格になるかどうかを考慮している。すなわち、「5人の犯人」といった場合には、「5人＝犯人」であるが、「5人の人生」といった場合には、「5人＝人生」とはならない。これは一種の「AのB」問題であるが、格助詞だけでなく、種々の要素が絡む可能性があり、現在検討中である。

今後は例文を増加し、数量詞の現れ方と取り扱い方についてさらに研究する予定である。

### 謝辞

数量詞に関する文献について、東京都立大学 荻野 綱男氏、東京大学大学院 飯田 朝子氏、電総研 橋田 浩一氏より有益な示唆をいただいた。ここに記して感謝の意を表する。

## 参考文献

- [1] 橋田浩一・郡司隆男：日本語の計量句と言語処理過程、日本認知科学会第9回大会発表論文集(1992)
- [2] Asako Iida: Modification of Numeral Quantifiers: “Feed *three* cats a fish” and “Feed a cat *three* fish”, *Colloquia* (Keio University) Vol.19 (1998)
- [3] 飯田 朝子：日本語「ひとつ」「ふたつ」等の表現における“一つ”の意味と用法－主要助数詞「個」「本」との比較を中心に－、東京大学言語学論集 17 (1998) pp.133-169
- [4] 池原 悟他編：日本語語彙大系、岩波書店(1997)
- [5] 中野 洋：「分類語彙表」形式による語彙分類表（増補版）、国立国語研究所(1996)
- [6] 落合 武：機械処理のための数量詞の研究、山形大学卒業論文(1999)
- [7] 荻野 綱男：現代日本語助数詞の意味変化の方向、筑波大学文芸・言語学系紀要、文藝言語研究 言語編 17 (1990)
- [8] 横山 晶一・加藤 貴子・廣重 拓司：格助詞「の」の分類と解析、言語処理学会第1回年次大会講演論文集 B1-2 (1995)