

# 英文科学技術文における高頻度名詞の分類について

1 R-3

河崎 裕司<sup>†</sup> 丸木 健次<sup>†</sup> 上福 潤<sup>‡</sup> 竹田 正幸<sup>†</sup> 松尾 文碩<sup>†</sup>  
<sup>†</sup>九州大学大学院システム情報科学研究科 <sup>‡</sup>九州大学工学部

## 1. まえがき

英文科学技術抄録文を論理式へ変換する第一段階として、原子論理式の述語記号に動詞を項に名詞句をそのまま単語列としてあてる方法が考えられる<sup>1)</sup>。この変換を行なうためには、文の動詞句の構造と名詞句の範囲を決定する必要があるが、この際発生する曖昧さを減らすために以下のことを考えた。

次の例文は、「... ルング・クッタ法により、解かれる。」という意味である。下線を施した単語 *method* は、文において手段を表していると考えられる。

In addition, the voltage, torque and speed equations are solved simultaneously by the Runge-Kutta method of the fourth order and the dynamic characteristics of the switched reluctance synchronous motor for variable control parameters are calculated and presented.

このようにして名詞自身によってその格を決定することができれば、原子論理式への変換を行なう際の曖昧さの解消に役立つ。

本稿では、高頻度名詞の用法を調査し、分類を行った。

## 2. 調査

まず、1969年から1993年までの25年分のINSPECテープの制御工学、計算機工学、情報工学分野における名詞の頻度を調査した。高頻度名詞の一部を表1に示す。調査対象を絞るために、これらの名詞のうち *author* などの特殊な名詞を除き、上位20名詞について

Classification of High Frequency Nouns in Scientific and Technical Documents

Yuji Kawasaki<sup>†</sup>, Kenji Maruki<sup>†</sup>, Jun Kamifuku<sup>†</sup>, Masayuki Takeda<sup>†</sup> and Fumihito Matsuo<sup>†</sup>

<sup>†</sup> Graduate School of Information Science and Electrical Engineering, Kyushu University, Hakozaki, Fukuoka 812-81, Japan

<sup>‡</sup> Faculty of Engineering, Kyushu University, Hakozaki, Fukuoka 812-81, Japan

表1 INSPECテープ制御工学、計算機工学、情報工学分野における高頻度上位名詞

単語	生起頻度	単語	生起頻度
system	712403	methods	117656
systems	380165	application	104399
data	347716	applications	101797
computer	300893	simulation	95046
method	233486	techniques	93541
information	214687	language	87772
software	170557	user	86500
algorithm	160337	management	82388
analysis	150894	algorithms	82038
authors	148142	solution	80329
problems	146009	technology	79554
network	135272	parameters	72010
performance	131581	knowledge	71723
author	127252	implementation	71489
development	124336	operaton	71444

て調査した。1993年のINSPECテープの制御工学、計算機工学、情報工学分野より、単文を抽出し、該当名詞を含む単文を調査対象とした。

次に、名詞毎にその名詞の用法を調査した。各名詞毎に20文ずつ実際に文を目で見て調査した。主語、目的語の核となる名詞は、それぞれ主語、目的語とした。他の名詞などを修飾している名詞は、修飾語とした。前置詞、成句などに係る名詞は、その他とした。結果を表2に示す。

## 3. 格の決定

表2において、修飾語については、格を決定できない。修飾語以外の名詞の用法について格の決定を行った。

以下に例文を示し、格を説明する。

表 2 用法調査結果

名詞	主語	目的語	その他	修飾語
system	5	6	4	5
data	1	5	2	12
computer	4	2	2	12
method	5	7	4	4
information	1	10	1	8
software	2	6	4	8
algorithm	9	5	2	4
analysis	5	8	5	2
network	6	1	4	9
performance	8	10	0	2
development	0	10	8	2
application	5	5	6	4
simulation	4	0	6	10
language	1	8	3	8
user	0	2	4	16
management	2	6	2	10
solution	5	8	5	2
technology	2	7	4	7
knowledge	3	5	1	11
implementation	1	13	4	2

● 手段格

Finally, these results are confirmed by computer simulation and a position measurement experiment of outdoor plants.

これは、「これらの結果は、... コンピュータシミュレーションによって確かめられた。」という意味であり、名詞 simulation は、文において手段を表していることから、そこでこれを手段格と呼ぶことにする。

● 被提示格

The authors propose a method to solve that problem.

The paper is concerned with the simulation of two-dimensional weakly correlated fields and their use in vehicle dynamics.

上の二つの文は、「著者らは、あの問題を解決する方法を提案する.」、「この論文は、... シミュレーションに関係する.」という意味であり、抄録文特有の主題を提示する文である。下線を施した名詞 method, simulation は、提示された主題を表している。そこでこ

れらを被提示格と呼ぶことにする。

● 所在格

Such a center line corresponds to the major axis of a two-dimensional point distribution in principal component analysis.

これは、「... 主成分解析における ... と一致する.」という意味であり、名詞 analysis は、所在を表している。そこでこれを所在格と呼ぶことにする。

● 主格

The method is based on zero isolation techniques for algebraic polynomials.

「この方法は、... に基づいている.」という意味である。名詞 method は、動作主と表している。そこでこれを主格と呼ぶことにする。

● 用途格

The modeling support system has a potential for application in environmental decision support systems and spatial decision support systems.

「... システムは、... アプリケーションのための可能性を持っている.」というこの文において前置詞 for は、用途・意図の意味で使われているので、この名詞 application は、用途格と呼ぶことにする。

以上のように、格を決定し、分類を行なった。

4. むすび

英文科学技術文の格構造化のために、名詞自身によってその格を決定することができるか調査した。高頻度名詞20単語について格を決定、分類し、考察を行なった。名詞 method, simulation, analysis などは、格を決定し易いようである。一方、名詞 data, user などは、比較的格の決定が困難である。

今後の課題として、調査対象を多くし、今回の格を検証することが挙げられる。

なお、本研究は、一部文部省科学研究費補助金 (# 07558162) の援助により行なった。

参考文献

1) 竹田, 松尾: 英文科学技術文における単文の原子論理式への変換, 情報処理学会第49回全国大会講演論文集 (1994).