

# 実験情報の分類に関する検討

## Examination of experimental Information classification

上田 和也  
Kazuya Ueda

岡山大学 太田研究室  
Ohta Laboratory, Okayama University

**概要** 本研究は、論文中の実験について書かれた段落をデータセットや実験結果などに分類し、特に段落を色分けして強調表示することで実験に関する内容の記載箇所を容易に把握できるようにすることを目的とする。本稿では、段落を分類するためにいくつかの手がかり語が利用可能か検討した。

### 1 はじめに

近年、さまざまな学術論文データベースの充実により、膨大な数の論文を手軽に入手できるようになった。しかし、論文は機械が処理できるデータとは異なり、人が読まなければ有効に活用できない。また、研究者にとって論文中の実験の手法や結果をまとめることは、研究内容の整理や比較に必要不可欠である。本研究では、論文中の実験に関連のある論文構成要素を実験情報と呼び、これらの情報を論文から自動抽出し、可視化することを目的とする。本稿ではとりわけ実験情報について書かれた段落を事前に決めた実験情報のクラスに分類する方法を検討する。

### 2 実験情報の抽出

本研究では平井らの作成したルールによる実験情報抽出プログラム[1]を利用する。この実験情報抽出プログラムは手がかり語リストを利用して実験情報の書かれた段落を抽出する。このプログラムで使用する段落の手がかり語の種類は以下の4種類である。

- FIGURE OR TABLE：図、表を意味する単語
- POINT：図、表に関連の深い単語
- MEASURE：実験の評価指標を表す単語
- EFFECT：精度の増加や減少等、効果を表す単語

具体的な手がかり語を表1に示す。“experiment”に対する“experimental”のように品詞等が異なる単語は、正規表現を利用して判別し、段落の手がかり語として利用する。実験情報の段落は表1の手がかり語と以下のルールを利用して抽出する。

- 一文中に FIGURE OR TABLE の単語と POINT の単語が同時に出現した場合その文を含む段落を実験情報とする。
- 一文中に MEASURE の単語と EFFECT の単語が同時に出現した場合その文を含む段落を実験情報とする。

表1：段落の手がかり語とその種類

種類	手がかり語
FIGURE_OR_TABLE	Figure, Table, Fig, Tab
POINT	show, list, draw, illustrate
MEASURE	F-measure, recall, precision, score, performance, experiment, result, dry-run, formal-run, method
EFFECT	improve, maximum, compare, average, point, degrade, best, better, submit, evaluation

### 3 実験情報の分類方法

平井らの実験情報抽出プログラムにより抽出された実験情報の書かれた段落を、手がかり語の出現を利用し“実験の評価方法”と“実験の結果”に分類する。

本稿ではこの分類のための手がかり語として result と evaluation, performance の適切性について実験により検討する。

### 4 実験情報の分類の実験

NTCIR-9[2]に投稿された論文5件を用いて、実験情報の段落の分類実験を行った。具体的には3つの手がかり語 result, performance, evaluation を利用して、実験情報の段落を“実験の評価方法”および“実験の結果”に分類できるか評価する。各手がかり語の出現する段落を“実験の評価方法”、“実験の結果”それぞれに分類した場合の適合率、再現率、F値を求め、手がかり語がどちらの分類に有用か調べる。そのため以下の式で分類の適合率、再現率、F値を求める。

$$\text{適合率} = \frac{\text{正しく分類された段落の数}}{\text{そのクラスに分類した段落の数}}$$

$$\text{再現率} = \frac{\text{正しく分類された段落の数}}{\text{そのクラスの段落の数}}$$

$$\text{F値} = \frac{(2 \times \text{適合率} \times \text{再現率})}{(\text{適合率} + \text{再現率})}$$

表 2, 3, 4 に, 手がかり語ごとの実験情報の段落の分類結果の適合率, 再現率, F 値をまとめる.

表 2: 手がかり語 result による分類結果

	適合率	再現率	F 値
実験の評価方法	0.029	0.125	0.007
実験の結果	0.500	0.680	0.576

表 3: 手がかり語 evaluation による分類結果

	適合率	再現率	F 値
実験の評価方法	0.500	0.375	0.429
実験の結果	0.125	0.040	0.061

表 4: 手がかり語 performance による分類結果

	適合率	再現率	F 値
実験の評価方法	0.167	0.250	0.200
実験の結果	0.467	0.280	0.350

実験の結果, 手がかり語 result は“実験の結果”の F 値が, 手がかり語 evaluation は“実験の評価の方法”の F 値が他の手がかり語よりも高く, またもう一方のクラスの F 値が低い結果となった. このため, この 2 クラスへの分類であれば誤る可能性が比較的 low, 手がかり語として利用できる可能性がある. ただし相対的に高いと述べた F 値も値自体は大きくないため分類の精度を上げるには他の手がかり語を併用するなど工夫が必要である.

一方手がかり語 performance は“実験の結果”の F 値は result より低く, “実験の評価の方法”の F 値は evaluation より低く, また 2 つの F 値にあまり差がないため, 手がかり語としては使用しにくいといえる.

## 5 まとめ

本研究では学術論文から抽出された実験情報の段落を, 手がかり語により内容ごとに分類する方法について検討した. また, 手がかり語ごとに実験情報の段落の分類結果を比較し, どの手がかり語がどの実験情報の判別に適切であるか実験により評価した. 今後の課題として, 実験情報の段落をどのように分類すべきか, 本稿の実験で使用しなかった他の手がかり語の利用可能性はどうか, 手がかり語が同時に複数種類出現したときにどのように分類すべきか, などが挙げられる.

## 参考文献

- [1] 平井久貴, 新妻弘崇, 太田学, 高須淳宏, “学術論文からの実験情報抽出の一手法”, DEIM Forum 2015, F3-1, 2015.
- [2] “NTCIR-9”, <http://research.nii.ac.jp/ntcir/ntcir-9/>.