

# 重要文抽出, 自由作成要約に対応した新聞記事要約システム YELLOW

大竹 清敬†      岡本 大吾      児玉 充      増山 繁  
kohtake@slt.atr.co.jp,    {okamoto, kodama}@smlab.tutkie.tut.ac.jp, masuyama@tutkie.tut.ac.jp  
豊橋技術科学大学 知識情報工学系

日本語の新聞記事を対象とした新聞記事要約システム YELLOW について報告する. YELLOW は, 「重要な情報を洩れなく抽出する」ことに重点をおいて作成した. 本システムは, 各文に対する重要度付与と文内要約のふたつの部分より構成されている. 重要度付与では, 主要語, 高頻度の名詞, 位置情報, 見解文であるか否かなど, 従来, 文の重要度を決定するにあたって重要であると言われてきた種々の情報を, 複合的に用いている. 文内要約では, 構文解析を用いて, 多重修飾語の削除, 換言処理, 例示の削除などを行う. 情報検索と自動要約の評価のためのワークショップ, NTCIR-2 に参加した結果, YELLOW は NTCIR-2 において, 全参加チーム中, 平均値で最も良い精度と評価された.

## A Summarization System YELLOW for Japanese Newspaper Articles

Kiyonori Ohtake†, Daigo Okamoto, Mitsuru Kodama, Shigeru Masuyama

Department of Knowledge-based Information Engineering  
Toyohashi University of Technology Toyohashi 441-8580, Japan

We propose a new automatic summarization system, YELLOW, for Japanese newspaper articles. YELLOW is designed to avoid omission of important information. The system was composed of two components, an extract-type summarizer and an abstract-type summarizer. In the extract-type summarizer, some features including main terms, high frequency words, location information in a paragraph, etc., are used to decide the weight of each sentence. The abstract-type summarizer summarizes sentences by deleting multiple modifiers for nouns and illustrations and by paraphrasing, etc. We participated in the tasks A-1 and A-2 of TSC in NTCIR-2 and the evaluation results show that YELLOW outperforms all other participants in average precision.

### 1 はじめに

計算機およびネットワークの発展によって, 膨大な量の文書が分散されて蓄積されるようになった. 我々は自らが抱える問題を解決するために, これらの文書から必要な情報を探しだし, 利用しなければならぬ. 一方で, 有史以来生物としての人間の情報処理能力はほとんど変化していない. そのため, 自動要約技術などにより, 読み手が読む文書の量を制御できることが求められている.

単一の文書に対する要約研究は, 長い歴史を持つ

ており [1], 重要な文を選ぶ, 重要文抽出型の要約や, 一文ごとに要約を行う文内要約などがある.

山本らによって開発された GREEN[2] では, 論説を対象に, 文書内の談話構造の利用による重要文抽出を行い, 連体修飾部の削除などの文内要約によって文章要約を試みている. しかしながら, GREEN は報道記事など, 論説以外の特徴をもつ記事への対応はしていない.

三上ら [6] は, TV ニュース原稿を対象として構文解析器を用いて文内要約を行っている. しかしながら, 連体修飾節の削除, 固有名詞へ係る修飾語句の削除のいずれにおいても, 重要部の認定が困難

† 現在 ATR 音声言語通信研究所

であると報告しており、重要な情報の削除が問題として指摘されている。

そこで、我々は、山本ら、三上らの問題点を考慮し、「重要な情報を洩れなく抽出する」こと、すなわち報知的 (informative) な要約を生成することに重点をおき、論説、報道記事の両方に対応した要約システム YELLOW<sup>1</sup> を開発した。また、国立情報学研究所主催の情報検索と自動要約評価のためのワークショップ NTCIR-2<sup>2</sup> の要約タスク TSC における A-1(重要文抽出)、A-2(自由要約作成) サブタスクに参加し、YELLOW の評価を行った。

YELLOW は、各文に対する重要度付与と各文の文内要約をもとに要約を行う。

重要度付与には、主要語、高頻度の名詞、位置情報、見解文であるか否かなど、従来、文の重要度を決定するにあたって重要であると言われてきた種々の情報 [3] を、複合的に用いている。重要度は、各情報におけるパラメータ値の総和により決定し、各パラメータ値は、NTCIR-2 の DRYRUN(事前に行われた評価の試行) に基づき、人手で決定した。なお、報道、論説の記事ごとに主要語の抽出箇所および各パラメータの重みを変えている。

文内要約では、KNP<sup>3</sup> による構文解析を行い、多重修飾に着目した削除、換言処理、直接引用表現の削除など、一文ごとの要約を行う。

換言処理では、山崎ら [4]、および若尾ら [5] の換言の手法を採り入れ、新聞記事を対象に、新たに冗長な文を簡潔に言い換えるテーブルを作成した。多重修飾に着目した削除では、三上ら [6] の手法のひとつである連体修飾節の削除を参考にして、「重要な情報の削除は極力行わない」という方針に基づき、新たに多重修飾における要約手法を構築した。さらに、直接引用表現を利用して要約知識の自動抽出を試みた児玉らの知見 [7] をもとに、要約知識を用いた直接引用表現の削除法を取り入れた。また、直接引用表現内の冒頭文の削除などヒューリスティクスに基づく要約手法も構築した。

以下、YELLOW の仕様、NTCIR-2 における評価を報告し、文内要約に使用した各手法の比較等の考察を述べる。

<sup>1</sup> YEt another summarization system with two moduLes using empiricaL knOWledge

<sup>2</sup> <http://research.nii.ac.jp/ntcir/workshop/work-ja.html>

<sup>3</sup> <http://www-lab25.kuee.kyoto-u.ac.jp/nl-resource/knp.html>

## 2 YELLOW の概要

YELLOW は、Linux Vine2.0 上で、Perl 言語を使用して作成した。システムは、記事の各文に対し、重要度付与を行う部分と文内要約を行う部分の二つから構成される。

YELLOW は、まず、主要語や語の頻度、段落内構造などの表面情報をもとに各文に重みを付け、各文の優先順位を決定する。次に、記事中の各文に対し、KNP による構文解析を行い、多重修飾語の削除、換言処理、例示の削除などの文内要約を行う。最後に、各文の優先順位と文内要約された各文をもとに、文を選択し、要約結果を出力する。

YELLOW の概要を、図 1 に示す。

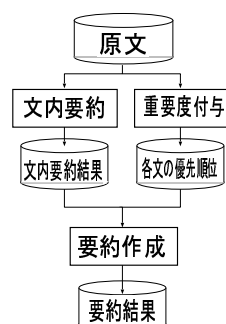


図 1: YELLOW の概要

## 3 重要度付与

記事中の各文に対し、五つのプロセスから重要度を付与する、その関係を図 2 に示す。

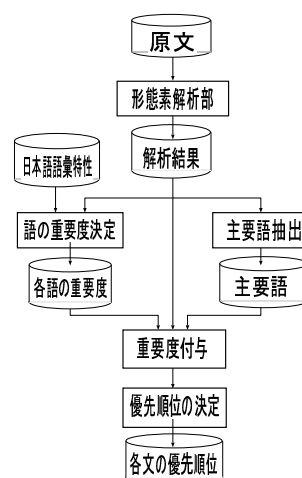


図 2: 重要度付与プロセスの概要

### 3.1 形態素解析部

形態素解析には、JUMAN<sup>4</sup>を用いる。解析した結果は、語の重要度決定、主要語抽出、重要度付与に使われる。

### 3.2 語の重要度決定

新聞記事において、記事中に多く出現する語は筆者が伝えたい情報であり、重要と考えられる。また、他の記事において多く出現する語は、一般的な語であり、重要ではないと考えられる。

そこで語の重要度決定部では、 $tf \cdot idf$ [12]の考え方に基づき、記事  $a$  における語  $w$  の重要度  $W(a, w)$  を以下の式により決定する。

$$W(a, w) = \frac{tf(a, w)}{\log(lp(w))} \quad (1)$$

ここで、 $tf(a, w)$ : 記事  $a$  での語  $w$  の頻度、 $lp(w)$ : 日本語語彙特性 [8] に記載されている  $w$  の頻度である。

式 (1) に示したように、 $idf$  を日本語語彙特性を用いて近似した。日本語語彙特性に記載されている単語の出現頻度は、朝日新聞の1985年から1998年の記事から計算されており、同じく新聞記事を対象としている本システムに適していたため、これを用いた。

### 3.3 主要語抽出

見出しは、記事の「究極的な要約」であり、その記事の主題を表すことが多い [2]。したがって、見出しに含まれる名詞は、少なからず主題に関係していると考えられる。また、記事中の重要な文も、見出しと同様に、主題について述べられている可能性が高い。報道記事では、冒頭段落に全体のまとめが書かれる傾向があり、冒頭段落に主題に関係する語が多く含まれていると考えられる。論説では、主要な結論が述べられる最終見解文と、前提を全く持っていない読者に初めて著者の持つ情報を提供する冒頭文に、主題に関係する語が多く含まれていると考える [2]。なお、見解文に関しては、3.4 節で定義する。

YELLOW では、見出し、および報道記事における冒頭段落中の文と、論説における冒頭文と最終見解文に含まれる名詞を主要語とする。

なお、YELLOW では、形態素解析によって未定義語のカタカナ、アルファベットと認定された語の

<sup>4</sup> <http://www-lab25.kuee.kyoto-u.ac.jp/nl-resource/juman.html>

ほとんどが、名詞であるという理由から、未定義語のカタカナ、アルファベットも名詞と同様の扱いをする。また、それぞれの主要語の主題への関連度は、語の重要度に比例すると仮定している。

### 3.4 重要度付与

各文の重要度は、主要語、語の重要度、記事構造、段落内構造、不要文、見解文か現象文かの計6つの要素に着目して決定する。重要度を求める際の各要素は、ヒューリスティクスによって決めた点数の加算によって計算する。

ただし、記事の種類により重要視される要素が異なるため、YELLOW では、論説と報道記事に対しそれぞれ異なった重み付けを行う。以下、6つの要素について説明する。

**主要語** 主要語は、その記事の主題など、重要な情報に関連しているとみなした語であり、それを含む文は重要であると考えられる。また、主要語がその文において主語(主格を表す格助詞をもつ語)である場合、その文は主題について述べていることが多い。そのため、より大きな重みをその文に付加する。

**記事中で頻度の高い語** 記事中に多く出現する語は筆者が伝えたい情報に関連しており、高頻度の語を含む文は重要であると仮定する。そこで、高頻度の語を含む文に重み付けをする。なお、YELLOW では、一記事に二回以上出現した名詞を、高頻度語と定義する。

**記事構造** 新聞記事において、冒頭文が重要であることは先に述べた。さらに、報道記事だけに着目すると、著者はより重要な情報ほど冒頭に書く傾向がある [3]。したがって、冒頭に近い段落ほど重要視する必要がある。また、報道記事の最終段落には、テーマに関する今後の展開など読者が関心をもつ情報が書かれる場合がある [2]。本手法では、このような情報も重要視し、最終段落も重み付けを行う。

**段落内構造** 各段落の冒頭文は、新しい主題に関する前提を持っていない読者に、著者の持つ情報を初めて伝達する役割をもつ。また、各段落の最終文は、著者がその段落の話を締めくくるといった特別な意図を持って書いた文と考えることができる。したがって、各段落の冒頭文および最終文は、その段落において重要な文と考えられる。

**不要文** 新聞記事には、段落のタイトルや補足説明など、不要な文が存在する。このような文は、記

事特有の記号(◇, =など)を含む傾向がある。また、括弧で囲まれた発言のみの文は、前後の文脈により情報を把握できることが多い。YELLOWでは、記事特有の記号を含む文、および発言内容のみを示す文は、不要文と定義し、重要視しない。

**見解文と現象文** 記事中の文は、著者の主張、意見、希望などを述べた文と、出来事、事実、現象を述べた文の二種類に大別することができる。以下、前者を「見解文」、後者を「現象文」と呼ぶ[2]。

論説においては、見解文が特に重要視される傾向がある。見解文を抽出するために、「～が必要である」「～すべきである」などの文末表現に注目する。YELLOWでは、GREENに基づき、あらかじめ作成した55個の文末パターンの表とのマッチングにより近似的に見解文の抽出を行う<sup>5</sup>。

各文の重要度は、表1に示す条件を独立に判定し、得点を総和することで計算する(ただし、不要文は、最終的に求めた文の重要度/10。報道記事の最終段落は、最終的に求めた文の重要度×10)。なお、主要度、高頻度語に関する得点は、文中に含まれる各語ごとに判定し、得点を総和する。

重み付けのパラメータの値は、DRYRUN(事前に行われた評価の試行)での評価結果に基づき、人手による試行錯誤で決定した。したがって、NTCIR-2の評価方法に特化した重み付けになっている。

表 1: 重要度付与条件とその得点

条件	得点	
主語が主要語である文	主要語の重要度×10	
主要語を含む文	主要語の重要度×2	
高頻度語を含む文	高頻度語の重要度×1	
各段落の冒頭文	20	
各段落の最終文	10	
不要文	文の重要度を1/10	
条件	報道記事限定	論説限定
第1段落	100	20
第2段落	50	0
第3段落	20	0
最終(まとめ)段落	文の重要度×10	0
見解文	0	10
それ以外	0	

### 3.5 優先順位の決定

各文の重要度をもとに、重要度の高い文から順番に、各文の優先順位を決定する。ただし、記事の冒頭文(冒頭文が不要文の場合は次の文)は、重要度にかかわらず、強制的に優先順位1番とする。

<sup>5</sup> [http://www.smlab.tutkie.tut.ac.jp/research/NL/DATA/yellow\\_data.html](http://www.smlab.tutkie.tut.ac.jp/research/NL/DATA/yellow_data.html)にて、全文末パターンを公開

## 4 文内要約

YELLOWで行われる文内要約は、以下の二つの基本方針に基づいて設計されている。

- 要約は、文ごとに独立に冗長な部分を削除(または換言)することで行う。文と文のまとめあげ、文間の冗長性の認定による削除など、複数の文を対象とした処理は行わない。
- 重要な情報が損なわれないことを重視し、それを削除することにより重要な情報が欠落するおそれのある冗長部の認定は避ける。

YELLOWでは、新聞記事の原文に対しKNPによる構文解析を行い、その結果を利用して、「補足情報の削除」、「直接引用表現の処理」、「多重修飾の削除」、「例示の削除」、「テーブルによる換言処理」の5つの手法により各文を文内要約する。

### 4.1 補足情報の削除

新聞記事において、丸括弧で囲まれた情報はフリガナや略称など、補足的な説明が多い。そこで、丸括弧で囲まれた情報および他の記号(=, <>など)で判断できる補足説明は削除する。

### 4.2 直接引用表現の処理

新聞記事内の直接引用表現は、基本的に重要度が低く、それが存在しない場合でも前後の文により理解できることが多い。また、第一文より、第二文以降の意見が重要視される傾向がある。そこで、直接引用表現が複数文で構成されるとき、第一文を削除する。

例: そのうえ「明日の公式試合には出なくてええ。背番号も返せ」と言われたという。→そのうえ「背番号も返せ」と言われたという。

ただし、第二文以降に第一文を参照する可能性のある指示詞が存在する場合は、削除を行わない。このような削除を行った場合、記事が理解しにくくなり、削除後の自然さも損なわれることを経験的に確認している。

また、新聞記事では、鈎括弧でくくられた直接引用表現の直後の節が、その引用表現の要約である場合がある[7]。YELLOWでは、直接引用表現を含む文の陳述表現が文献[7]に示されるパターンにマッチする場合、発言部分を削除しても意味内容を保持できると仮定し、直接引用表現部分を全て削除する。

例：検察側は「捜査段階で事実を認めていた」と主張して、タイミングを計って証拠申請をする構え。  
→検察側はタイミングを計って証拠申請をする構え。

### 4.3 多重修飾の削除

ひとつの語に対して文節が二つ以上修飾している状態を「多重修飾」、特に二つの文節が修飾している状態を「二重修飾」と定義する。本章では、以後、新聞記事に現れる多重修飾のほとんどが二重修飾であることから、二重修飾のみを扱う。

YELLOW では、名詞に係る二重修飾があった場合、その片方の文節を削除しても意味は大きく変化しないことが多いと想定し、一方の修飾要素を削除する。なお、YELLOW では、基本的に名詞を限定する働きの強さを基準として、どちらの修飾要素を削除するかを決定する。また、名詞を限定する働きが同程度の場合は、前方の修飾要素を削除する。

例：政党が軽い存在となり、政党に対する慢性的な不信が渦巻いている。→政党が軽い存在となり、政党に対する不信が渦巻いている。

また、多重修飾の削除において、削除した結果が不自然になると考えられる場合は、削除を行わない。この削除の回避は、二重修飾の修飾要素と被修飾要素との間に制約を設けることにより実現する。二重修飾の前方と後方の修飾成分ならびに被修飾要素の名詞を含む文節が与えられたときの削除規則の例を表2に示す。なお、表中の“—”は、特に制約がないことを意味する。このような規則が合計36個存在する<sup>6</sup>。

表 2: 二重修飾の削除規則例

前方	後方	被修飾要素	動作
～の	～の	—	削除しない
—	—	～との	削除しない
～という	～の	～の	後方を削除
連体修飾節	形容詞	—	後方を削除
—	～な	こと…	削除しない

YELLOW は、係り受け関係を調べる際に用いる KNP が、解析を誤る可能性を想定し、いくつかの対処を行っている。

**主題部内から外への係り受け** 主題部から外への係り受け関係を認めず、削除を行わない。これは、主題部内から外への係り受け関係を持つ文は、その距離が大きくなる傾向があり、KNP の解析誤り

<sup>6</sup> [http://www.smlab.tutkie.tut.ac.jp/research/NL/DATA/yellow\\_data.html](http://www.smlab.tutkie.tut.ac.jp/research/NL/DATA/yellow_data.html)にて、全削除規則を公開

が起りやすいためである。YELLOW では、文頭から、取り立て助詞を末尾にもつ文節までを「主題部」と定義する。また、取り立て助詞のあとに格助詞が続く場合もこれに該当する。

**多重修飾の可能性のある構造の処理** 多重修飾にも係わらず、KNP の解析誤りにより、正しく多重修飾と解析されない場合がある。YELLOW では、「連体修飾節+名詞の+名詞」という構造の文においては、KNP の構文解析結果にかかわらず、多重修飾とみなし、連体修飾節を削除する。

ただし、文の大意を保つため「名詞の+名詞」という構造のみでは被修飾名詞の内容が十分に伝わらない場合、連体修飾節を削除しない。「名詞の+名詞」という構造のみで被修飾名詞の内容が十分に伝わるかどうかは、「名詞の」に含まれる名詞が抽象的すぎないかどうかで判断する。なお、YELLOW では、日本語語彙体系 [9] の各カテゴリを代表する名詞を抽象的な語とする。

### 4.4 例示の削除

「～などで」「～などの」のような例示は、広い意味での修飾と考えられ、削除しても意味的に変化が生じないと近似的に仮定する。YELLOW では、「～などの+名詞」という表現において「～などの」を削除する。ただし、「～などの」に係る名詞が抽象的すぎる場合は削除しない。また、「～などで」が用言に単独に係る場合、「～などで」を削除する。

例：経済や外交戦略などの専門知識はもとより、→専門知識はもとより、

### 4.5 テーブルによる換言処理

新聞記事に見られる語尾のまわりくどい表現は、体言止めにするなどの換言により要約できる。また、前文との結束性を表す冒頭の接続詞は、前の文が抽出されなければ意味をもたないため、冒頭の接続詞を削除する必要がある。

YELLOW では、このような処理を文字列照合に基づいて、ある文字列を別の文字列へ変換する規則(1対1換言)によって実現する。この変換規則の集合を換言テーブルと呼ぶ。このような規則を、文末表現の簡潔化を中心に、人手で106用意した<sup>7</sup>。

例1：…決まらないようだ。 →…決まらない。

例2：そんな中、…。 →…。

<sup>7</sup> [http://www.smlab.tutkie.tut.ac.jp/research/NL/DATA/yellow\\_data.html](http://www.smlab.tutkie.tut.ac.jp/research/NL/DATA/yellow_data.html)にて、全換言テーブルを公開

## 5 要約作成

YELLOW は、文内要約で要約された各文と、重要度付与による各文の優先順位をもとに、使用する文を選択し、それを要約とする。このとき、規定文字数を入力することで、YELLOW は、規定文字数を越えないように要約を行う。

要約結果を出力する際のアルゴリズムを以下に示す。なお、要約文として文を使用することを「採用する」と表現する。

### 要約作成アルゴリズム

1. 各文に対してテーブルによる換言処理を除く文内要約を行う。
2. 規定された文字数を越えるまで、優先順位が大きい順に文を選ぶ。
3. 全ての文に対してもう一度、テーブルによる換言処理を除く文内要約を行う。
4. 選択した文全ての長さが、規定された文字数以下ならば要約結果を出力して終了。
5. 選択した文に対してテーブルによる換言処理を行う。
6. 選択した文全ての長さが、規定された文字数以下ならば要約結果を出力して終了。
7. 選択した文の中で最も優先順位の低い文を除き、規定文字数-現在の文字数以下の長さの文をまだ選択されていない文のなかから重要度が高い順に探す。
8. 該当する文があれば、それを採用し、テーブルによる換言処理を行い、要約結果を出力して終了。
9. 該当する文が存在しない場合、8. の処理を終えた段階で選択されている文を要約として出力する。

## 6 評価

YELLOW の評価を行うため、NTCIR-2 の要約に関するタスクのうち、A-1 サブタスク、A-2 サブタスクに参加した。

重要文抽出の評価を行う A-1 サブタスクでは、YELLOW の優先順位をもとに文を出力することで、文への重要度付与に対する精度を評価した。

また、自由作成要約の評価を行う A-2 サブタスクにおいて、YELLOW による要約の有効性を評価した。

### 6.1 A-1(重要文抽出) サブタスクの評価

A-1 タスクでは、人間が選択した重要文とシステムが選択した重要文との間の一致度を元に評価が行われた。評価尺度としては、以下の三つを用いた。

$$\text{再現率} = \frac{\text{システムが選んだ文のうちで正解文の数}}{\text{人間が選んだ正解文の総数}}$$

$$\text{精度} = \frac{\text{システムが選んだ文のうちで正解文の数}}{\text{システムが選んだ文の総数}}$$

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

要約率 10%、30%、50%ごとに、30 記事 (報道 15 記事、論説 15 記事) の平均を求めたときの F 値、および、記事を報道限定、論説限定にしたときの結果を表 3 に示す。

表 3: A-1 タスクの評価結果

	F 値	報道の F 値	論説の F 値
10%	38.6	47.8	29.4
30%	46.8	49.2	49.2
50%	62.2	63.4	60.9
ave	49.2	53.4	44.9

#### 6.1.1 考察

結果より、報道記事に比べ、論説の重要文抽出の精度が全体的に低いが、これは、論説が報道のように決まった構成を持たないことや、報道に比べ記事中の文数が多い等の特徴が影響しているためである。また、10%への要約のように、少数の文を選択する場合、論説においては、冒頭文強制採用の誤りによる影響が大きく現れた。これは、冒頭文が必ずしも全体の概要にはならず、筆者によって書き出しが大きく変わるためであり、冒頭文が全体の概要を示しているかを判定した上で、採用する等の改良が必要である。

#### 6.1.2 重要度付与に用いた各要素の比較

YELLOW では、6つの要素をもとに各文の重要度を付与している。重要度の決定に用いた各要素の効果を、それぞれ単独に無効化することで、比較した結果を表 4 に示す。表中のそれぞれの値は、10%、30%、50%の平均である。

各要素を比較した結果、全体的には不要文の影響が一番大きいことがわかる。つまり、精度をあげるためには、重要ではない情報をいかに丁寧に取り除くかが基本となる。

表 4: 重要度付与に用いた各要素の比較

無効化した要素	F 値	報道の F 値	論説の F 値
見解文	48.4(-1.6%)	53.4( 0%)	43.4(-3.3%)
記事構造	48.0(-2.4%)	50.9(-4.7%)	45.1(+0.4%)
段落内構造	46.4(-5.7%)	52.3(-2.1%)	40.5(-9.7%)
主要語	49.4(+0.4%)	53.2(-0.4%)	45.7(+1.8%)
記事中の高頻度語	47.8(-2.8%)	53.3(-0.2%)	42.3(-5.8%)
不要文	44.5(-9.6%)	50.0(-6.4%)	39.0(-13.1%)

また、段落内構造による重み付けは、大きな影響があり、特に、段落冒頭文の重み付けのみを無効化した場合、F 値は、 $-7.7\%$ と大きく精度が落ちた。これは、文章をいくつかのまとまりに分けた場合、筆者はその冒頭に重要な情報を書く傾向が強いためである。したがって、段落のみならず、筆者の考えたまとまりと同様に文章を分けることが、より精度をあげるための有効な手段と考える。

更に、報道記事においては、記事構造による重み付けが大きく影響しており、冒頭に近い段落ほど重要な構造を持っていることが確認された。また、論説においては、見解文や記事中の高頻度語が重要であることがわかった。しかしながら、論説において、見出しなどから抽出した主要語は、効果を発揮しなかった。これは、主要語を主語とする文が必ずしも重要ではなく、むしろ主題に対する意見が述べられる、それ以降の文を重要視すべきことを示唆している。したがって、主要語と各文の重要度の関係を再検討する必要がある。

## 6.2 A-2(自由作成要約) サブタスクの評価

A-2 タスクでは、主観評価と content-based での評価という二種類の評価が行われた<sup>8</sup>。

主観評価では、記事ごとに「テキストとして読みやすいかどうか」と「元テキストの重要な内容を不足なく記述しているかどうか」の二つの観点から、要約筆記の専門家が要約を評価した。なお、評価値は、読みやすいものから、1, 2, 3, 4 とし、同様に内容の点で見て、良いものから、1, 2, 3, 4 としている。よって、評価値の低いほうが良い結果となる。

content-based での評価は、まず、人間の作成した要約およびシステムの作成した要約を共に形態素解析し、内容語のみを抽出する。そして、人間の

作成した正解要約の単語頻度ベクトルとシステムの要約の単語頻度ベクトルの間のコサイン距離を計算し、どの程度内容が単語ベースで類似しているかという値を求める。また、正解要約は、人間が自由作成した要約(以下、FREE)と、人間が重要個所抽出により作成した要約(以下、PART)の二種類がある。それぞれの正解要約に対し、要約率 20%と 40%になるように規定文字数が指定され、評価された。評価結果を、表 5, 表 6 に示す。

表 5: 主観評価の結果

	評価値 (全参加システムの平均)
読み易さ 20%	2.53(3.16)
内容評価 20%	2.93(3.24)
読み易さ 40%	2.73(3.05)
内容評価 40%	2.77(3.12)

表 6: content-based での評価結果

	評価値 (全参加システムの平均)
FREE 20%	0.4727(0.4418)
FREE 40%	0.6483(0.6065)
PART 20%	0.5137(0.4740)
PART 40%	0.6608(0.6342)

### 6.2.1 考察

今回、重要度付与によって採用した重要文(22812 文字)に対して、文内要約で行った冗長部の削除(2129 文字)の要約率は約 91%である。YELLOW では、重要な情報の欠落を防ぐため、大胆な削除は行わず慎重な要約を行った。その方針は比較的成功し、良好な評価にも結び付いているといえる。

### 6.2.2 各手法の比較

YELLOW では、5つの手法を併用することにより文内要約を行っている。それぞれの手法単独による削除文字数、content-based での評価を調査した結果、表 7 のようになった。

<sup>8</sup> <http://oku-gw.pi.titech.ac.jp/tsc/>

表 7: 文内要約における手法の比較

手法	削除文字数	割合 (%)	使用回数	平均削除 文字数	content-based の評価	
					FREE 20%	FREE 40%
補足情報の削除	662	31.1	335	1.97	0.4631	0.6067
直接引用表現の処理	298	14.0	8	37.25	0.4679	0.6198
多重修飾の削除	729	34.2	61	11.95	0.4489	0.6199
例示の削除	126	5.9	6	21.00	0.4627	0.6326
テーブルによる換言処理	314	17.1	89	3.53	0.4652	0.6206

各手法のうち最も要約に貢献した手法は、一度に削る文字数が多く、使われる場面も多い、多重修飾の削除であった。一方で、content-based での評価は、他の手法に比べ、やや精度が落ちている。これは、必要な情報の削りすぎがあるためであり、特に「連体修飾節+名詞の+名詞」という構成の連体修飾節を削除した場合に、不自然に感じられる場合が多いとの印象を受けた。この点に関して、厳密に評価し、改良することは今後の課題である。

テーブルによる換言処理も、一度に削る文字数は少ないものの、多くの文に適用できるため、要約において有効な手法といえる。ただ、換言表現を行うためのテーブルを手で作成しているため、より多くの文に適用させるためには、テーブルを拡張する必要がある。

## 7 まとめ

報道、論説の両方に対応した要約システムを紹介した。このシステムは、ヒューリスティクスに基づく、各文の重要度付与と、文内要約により、要約を行うものであった。また、文内要約では、GREENをはじめ様々な先行研究[3]で提案された手法を改良して取り入れ、新たに多重修飾の削除手法などを考案し構築した。

本システムの評価により、重要度付与に関しては、報道記事に比べ論説に対する精度が低いことがわかった。報道記事は、冒頭に近いほど重要視されるような定型的な構成が多いのに比べ、筆者により大きく構成の変わる論説は、対応しにくく抽出が難しいのが現状である。

文内要約では、重要な情報の欠落を防ぐため、大胆な削除は行わず慎重な要約を行うという方針により、良好な結果を得られた。

今後の課題としては、YELLOW では考慮しなかった談話構造による文の結束性をどう取り入れ

るか、または、複数文のまとめあげ等による、自然さの向上が望まれる。

## 8 謝辞

本研究で使用したコーパスは、毎日新聞 CD-ROM'94~98 版から得ており、その使用許可を頂いた同社に深謝する。また、日本語語彙大系から意味分類を取得するために用いた形態素解析システム ALTJAWS ver.2.0. の使用許可を頂いた(株)日本電信電話に深謝する。

## 参考文献

- [1] Mani, I. and Maybury, M. T. eds.: *Advances in Automatic Text Summarization*, MIT press(1999)
- [2] 山本和英, 増山繁, 内藤昭三: 文章内構造を複合的に利用した論説文要約システム GREEN, 自然言語処理, Vol.2, No.1, pp.39-55 (1995).
- [3] 奥村学, 難波英嗣: テキスト自動要約に関する研究動向, 自然言語処理, Vol.6, No.6, pp.1-26 (1999).
- [4] 山崎邦子, 三上真, 増山繁, 中山聖一: 聴覚障害者用字幕生成のための言い換えによるニュース文要約, 言語処理学会 第 4 回 年次大会論文集, pp.646-649 (1998).
- [5] 若尾孝博, 江原暉将, 白井克彦: テレビニュース番組の字幕に見られる要約の手法, 情報処理学会研究報告書 NL-122-13, pp.83-89(1997)
- [6] 三上真, 増山繁, 中川聖一: ニュース番組における字幕生成のための文内短縮による要約, 言語処理学会論文誌「自然言語処理」, Vol.6, No.6, pp.65-81 (1999).
- [7] 児玉充, 片岡明, 増山繁, 山本和英: 直接引用表現を利用した要約知識の自動抽出の試み, 言語処理学会 第 6 回 年次大会論文集, pp.241-244(2000).
- [8] NTT コミュニケーション科学基礎研究所: 日本語語彙特性, 三省堂(2000).
- [9] NTT コミュニケーション科学基礎研究所: 日本語語彙大系, 岩波書店(1999).
- [10] 大野晋, 浜西正人: 角川類語新辞典, 角川書店(1981).
- [11] 野田尚史: 語順を決める要素, 言語, Vol.29, No.9, pp.22-27(2000)
- [12] Gerard Salton: "Automatic Text Processing", Addison-Wesley(1998).