

テキスト自動要約に関する研究動向 (巻頭言に代えて)

奥村 学[†] 難波 英嗣[†]

本稿では、これまでの (主に領域に依存しない) テキスト自動要約手法を概観する。特に重要箇所の抽出を中心に解説する。また、これまでの手法の問題点を上げるとともに、最近自動要約に関する研究で注目を集めつつある、いくつかのトピックについてもふれる。

キーワード: テキスト自動要約, 重要文抽出, テキスト構造, 結束性, 機械学習, テキストのジャンル, 走り読み, ハイパーテキスト

Automated Text Summarization: A Survey

MANABU OKUMURA[†] and HIDETSUGU NANBA[†]

In this article, we try to survey the state of the art of (mainly domain-independent) automated text summarization techniques. We also try to provide readers with the limitations of current technology, and some new trends in the research field.

KeyWords: *automated text summarization, sentence extraction, text structure, cohesion, machine learning, genre, skimming, hypertext*

1 はじめに

電子化されたテキストが世の中に満ち溢れ、情報洪水という言葉が使われるようになってからかなりの歳月を経ている。しかし、残念ながら、我々の情報処理能力は、たとえ処理しなければならぬ情報が増えたとしても、それほど向上はしない。そのため、自動要約技術などにより、読み手が読むテキストの量を制御できることが求められている。また、近年情報検索システムを利用する機会も増えているが、システムの精度の現状を考慮すると、ユーザは、システムの提示した候補が適切なものであるかどうかをテキストを見て判断せざるを得ない。このような場合、要約をユーザに提示し、それを見て判断を求めるようにすると、ユーザの負荷を減らす支援が行なえる。

自然言語処理の分野では、近年頑健な解析手法の開発が進み、これと、上に述べたような、自動要約技術の必要性の増大が重なり、自動要約に関連した研究は、90年代の中頃になって、再び脚光を集め始めている。市販ソフトウェアも続々と発売されており、アメリカでは DARPA 支援の Tipster プロジェクトで要約が新しい研究課題とされたり (Hand 1997)、また、ACL、AAAIなどで要約に関するワークショップ、シンポジウムが相次いで開催され、盛況で活発な議

[†] 北陸先端科学技術大学院大学 情報科学研究科, School of Information Science, Japan Advanced Institute of Science and Technology

論が交わされた。日本でも、98年3月の言語処理学会年次大会に併設して、要約に関するワークショップが開催され、それを機会に本特集号の編集が企画された。

本稿では、このような現状を鑑み、これまでの(主に領域に依存しない)テキスト自動要約手法を概観する。また、これまでの手法の問題点を上げるとともに、最近自動要約に関する研究で注目を集めつつある、いくつかのトピックについてもふれる。本特集号の各論文が、テキスト自動要約研究として、どのような位置付けにあるかを知る上で、本稿が参考になれば幸いである¹。

要約研究は時に、情報抽出(Information Extraction)研究と対で(あるいは、対比して)述べられることがある。どちらも、テキスト中の重要な情報を抜き出すという点では共通するが、情報抽出は、あらかじめ決められた「枠」を埋める形で必要な情報を抜き出す。そのため、領域に依存してあらかじめ枠を用意する必要があったり、また、領域に依存したテキストの特徴を利用した抽出手法を用いたりするため、領域を限定することが不可欠となる²。

要約は、原文の大意を保持したまま、テキストの長さ、複雑さを減らす処理とも言えるが、その過程は、大きく次の3つのステップに分けられるとされる: テキストの解釈(文の解析とテキストの解析結果の生成)、テキスト解析結果の、要約の内部表現への変形(解析結果中の重要部分の抽出)、要約の内部表現の要約文としての生成。しかし、これまでの研究では、これらのステップは、テキスト中の重要箇所(段落、文、節、など)の抽出およびその連結による生成として実現されることが多かった。そのため、本稿では以後重要箇所の抽出を中心に解説する。

2節では、まず重要箇所抽出に基づく要約手法について述べる。2.1節で重要箇所抽出に用いられてきた、さまざまな情報を取り上げ、それぞれを用いた要約手法について述べる。2.2節では、それらの情報を統合して用いることで、重要箇所を抽出する研究について概観する。2.3節では、重要箇所抽出に基づく要約手法の問題点について述べる。

このようなテキスト要約手法が伝統的に研究されてきた一方で、近年要約を研究するに当たって考慮すべき要因として、以下の3つが提示されている(Sparck Jones 1999 to appear)。

- (1) 入力 の 性質-テキストの長さ、ジャンル、分野、単一/複数テキストのどちらであるか、など
- (2) 要約の目的-どういう人が(ユーザはどういう人か)、どういう風に(要約の利用目的は何か)³、など
- (3) 出力の仕方

1 各論文の個別の紹介は、増山氏の編集後記を参照して頂きたい。

2 情報抽出研究に関する解説としては、(Cowie and Lehnert 1996; 関根 1999)を参照されたい。また、DARPAが支援する情報抽出のプロジェクトであるMUC(Message Understanding Conference)に関しては、若尾の解説(若尾 1996)を参照して頂きたい。

3 要約は一般に、その利用目的に応じて、次の2つのタイプに分けられることが多い(Hand 1997)。

indicative: 原文の適切性を判断するなど、原文を参照する前の段階で用いる

informative: 原文の代わりとして用いる

たとえば、入力テキストのジャンルによっては、重要箇所抽出による要約が難しいものも考えられるし、また、要約というものの自体が考えにくいものもあり得る。ユーザの持つ予備知識の程度に応じて、要約に含める情報量は変えるべきであると考えられるし、また、利用目的が異なれば、その目的に応じた適切な要約が必要と考えられる。

これまでの伝統的な要約研究は、このような要因に関して十分な考慮をしたものとは必ずしも言えない。しかし、これらの要因を考慮して、入力の性質、要約の目的に応じた適切な要約手法を開発する動きが活発になってきている。このような、自動要約に関する研究で最近注目を集めつつある、いくつかのトピックについても本稿ではふれる。

3, 4, 5 節ではそれぞれ、抽象化、言い換えによる要約、ユーザに適応した要約、複数テキストを対象にした要約に言及する。6, 7 節ではそれぞれ、文中の重要箇所抽出による要約、要約の表示方法について述べる。8 節では、要約の評価方法について説明する。

2 重要文抽出による要約手法

2.1 重要文抽出に用いられるテキスト中の特徴について

1950年代まで歴史を遡ることができるテキスト自動要約研究のこれまでの多くのものは、テキスト中の文(あるいは、形式段落)を1つの単位とし、それらに何らかの情報を基に重要度を付与し、その重要度で順序付け、重要な文(形式段落)を選択し、それらを寄せ集めることで、要約を作成する。本節では、この重要度評価の際に用いられている、テキスト中の(主に表層的な)情報について述べる。

Paice(Paice 1990)はこの情報を7つに分類しているが、ここではそれも参考にした上で、以下の7つの情報を取り上げ、各小節で説明する。

- (1) テキスト中のキーワードの出現頻度を利用する、
- (2) テキスト中あるいは段落中での位置情報を利用する、
- (3) テキストのタイトル等の情報を利用する、
- (4) テキスト中の文間の関係を解析したテキスト構造を利用する、
- (5) テキスト中の手がかり表現を利用する、
- (6) テキスト中の文あるいは単語間のつながりの情報を利用する、
- (7) テキスト中の文間の類似性の情報を利用する

(1) テキスト中の単語の出現頻度の利用

テキスト中によく出現する内容語はテキストの主題を示す傾向があるとの仮定が情報検索分野などではしばしば用いられる。この仮定に基づき、テキスト中で出現頻度の高い名詞をキーワードと考えたり(tf法)、また、これに合わせて、出現するテキスト数も考慮することで、そ

のテキスト固有の出現の度合を計算したり (tf*idf 法) など、情報検索分野では、さまざまな単語の重み付け技法が用いられている (Salton 1989).

テキスト中の出現頻度に基づき単語に重要度を与えるという、このような考え方を利用し、単語の重要度を基に、文に重要度を付与するという重要文抽出手法が、自動要約研究の開始当初である 1950 年代から提案されている (Luhn 1958; Edmundson 1969; Zechner 1996; 若尾, 江原, 村木, 白井 1997). 単語の重要度から文の重要度を計算する手法はさまざま提案されているが、その一例としては、文中に出現する単語の重要度の総和を文の重要度とするものがある.

Zechner (Zechner 1996) は、単語を tf*idf 法で重み付けし、文中に出現する単語の重みの総和を文の重要度とする重要文抽出手法の評価を行なっている. 人間の被験者の要約と比較し、再現率/適合率を計算した結果、人間の被験者同士の比較による精度と大差ない結果を得ている. また、新聞記事を対象としているので、先頭数文を抽出する手法との比較を行っており、tf*idf 法を用いた手法の方が良い結果を得ているとしている.

また、単語ではなく、テキスト中で隣接する単語の対の頻度を基に、文に重要度を付与する手法を鈴木らは提案している (鈴木, 栃内 1988). Aone ら (Aone, Okurowski, Gorlinsky, and Larsen 1997) は、複合語を自動的に抽出し、それらの頻度も考慮する手法を提案している.

(2) テキスト中での位置情報の利用

テキストは、ジャンルに依存して、ある程度構造に規則性が有ると通常考えられている. たとえば、学術論文は、序論、本論、結論のような構造を持つし、新聞は、見出し、小見出しの後に、本文が来ることが多い. このような、ジャンルにより決まったテキストの構造を重要箇所抽出に利用する研究を本節と次節では紹介する.

テキストの構造から、テキスト中での重要な箇所の位置はある程度予測可能であると仮定して、テキスト中での文の位置情報をその文の重要度計算に利用する手法がいくつか考えられている. 論説文の場合に、テキスト全体のまとめは書き出しや結び近くにあると仮定するものや、重要な文はテキストの先頭、最後、段落の先頭、最後、節の見出しの直後にあると考える (Edmundson 1969) ものなどはその一例と言える. また、新聞記事を対象とした重要文抽出では、本文の先頭数文を抽出するのが良いとされる (この手法は lead 手法と呼ばれることが多い) (Brandow, Mitze, and Rau 1995) のも、新聞記事の構造 (本文中では大意をまず先頭に示す) に基づいた、位置情報を利用した手法とすることができる.

Brandow ら (Brandow et al. 1995) は、新聞、雑誌の記事を対象に、60, 150, 250 語の長さの要約文において、lead 手法と、「単語の出現頻度」、「見出しの情報」などを利用した、彼らのシステム ANES の出力を、受容可能性 (acceptability) で比較した結果、92%対 74%の差で lead 手法の方が良かったと報告している. 受容可能性は、経験あるニュース分析者が、原文と照らし合わせて、要約の内容と読み易さに関して判断する指標である.

Wasson(Wasson 1998) は, Brandow らの実験を追試し, lead 手法が有効である (あるいは, 有効でない) テキストはどのような種類かということを確認するため, より詳細な評価を行っている. Brandow らと同様, ニュース分析者が受容可能/不可能の判定を要約に対して行なう評価法を用い, テキストの種類 (ジャンル), 分野, 長さなどの違いにより, lead 手法の評価がどのように変化するかを調査している. Lead 手法は, テキストのジャンルに関して, 評価にばらつきが見られ, ニュースで高い受容可能性を示したが, その細分類の中では, Review で低い受容可能性を示した. 新聞記事に限定した場合の受容可能性は 95.5%であった.

(3) テキストのタイトル等の情報の利用

ジャンルにより決まったテキストの構造から得られる, もう一つの情報として, 本文以外に, テキスト中に付与されたタイトル, 見出しの情報がある. たとえば, 学術論文の場合は, テキスト自体がタイトルを持つ場合もあり, また, 各章, 節にもタイトルが付与されることが多い. また, 新聞には, 見出し (headings), 小見出しが本文とは別に付与されることもある.

このタイトル, 見出しは, ある意味で, テキスト本文の非常に簡潔な要約とも考えられる. そのため, タイトル, 見出しに現れる内容語を含む文が重要であると考え, タイトル, 見出し中の単語を重要文抽出に利用する手法がいくつか提案されている.

(Edmundson 1969) を始めとして, 最近では, 見出しに含まれる名詞を多く含む文を重要として抽出する (亀田 1996; 仲尾 1997; Ochitani, Nakao, and Nishino 1997) などその一例と考えられる⁴.

(4) 文間の関係を解析したテキスト構造の利用

自然言語処理の分野では, テキスト中の接続詞等の手がかり語情報などを基に, 文間の構造を解析し, テキスト構造を得る研究がいくつか見られる (たとえば, (福本, 安原 1991; Sumita, Ono, Chino, Ukita, and Amano 1992; Kurohashi and Nagao 1994)). このようにして得られたテキスト構造を利用して重要文を抽出する研究が近年見られるようになってきている (Miike, Itoh, Ono, and Sumita 1994; Marcu 1997).

Marcu(Marcu 1997) は, 修辞構造解析の結果得られる核 (nucleus) がテキスト中の重要箇所検出に有効であるかどうかを実証するため, 核と重要箇所間に相関関係があるかどうかを示す実験を行なっている. 5 テキストを 13 人の被験者に提示し, 3 段階評価で重要文の抽出を行なうと同時に, 2 人の計算言語学者がテキストを修辞構造解析し, 構造木を生成した. そして, その構造木を基にそれぞれの文に重要度を付与した. 結果として, 被験者の作成した要約との比較で, 計算言語学者の修辞構造を用いた要約は, 再現率 67.5%, 適合率 78.5%を得てい

⁴ (仲尾 1997) は, 見出し中の同じ名詞に関連する文を複数抽出しても冗長であるという考えに基づき, 単純に文の重要度の順に文を選択せず, 独自の選択手法を提案している.

る。Marcu はまた、450 種類の discourse marker を用いた修辞構造解析器を作成し、それを用いた要約作成実験も試みている。結果として、被験者の作成した要約との比較で、修辞構造解析器を用いた要約が再現率 66%、適合率 68%を得ている。

日本語に対しては、Miike らの研究 (Miike et al. 1994) がある。接続詞、照応表現などの手がかりを用いた規則集合により、文間の関係を解析し、テキスト構造を抽出するシステムを作成し、得られたテキスト構造に基づき、文に重要度を付与し、要約を作成している。人間の作成したテキスト構造と解析により得られたテキスト構造の比較、抽出した重要文と被験者の抽出した重要文の比較、抽出した重要文を用いて、検索テキストの適切性判断を被験者に求めた際の、所要時間、判断の精度 (再現率/適合率) により、テキスト構造解析器およびそれを用いた要約手法の評価を行なっている。

解析により得られたテキスト構造を利用して重要文を抽出する手法の利点としては、

- 長さに応じた要約を、得られた構造木のそれぞれのレベルで作成できる、
- テキスト構造に基づいて重要文を抽出しているため、単語の頻度などを用いた手法に比べ、首尾一貫性の高い要約が作成できる可能性がある

点があげられる。

(5) 手がかり表現の利用

(1) で述べたような、テキスト、文の主題を表す内容語ではないが、テキスト中の重要箇所を指示すると考えられる手がかり表現がいくつか存在する。たとえば、学術論文などでは、‘this report’, ‘in conclusion’, ‘our work’ などの表現は、論文の主題を表す文中に出現すると考えられる。このような手がかり表現を利用して、テキスト中の重要文を抽出する研究も存在する (Edmundson 1969)。これとは逆に、重要文と負の相関関係にあると考えられる手がかり語を考慮することもできる。たとえば「」などの例示を示す接続語で始まる文は重要度が低いと考えられるのはその一例である。

(6) 文間、単語間のつながりの利用

本節と次節では、テキスト中の文間のつながりの情報を重要文抽出に利用する手法について説明する。

Skorokhod'ko (Skorokhod'ko 1972) は、文をノード、文間の関係をリンクとするグラフでテキストを表現し、多くの文と関係のある文が重要であるという考えに基づき、重要文を抽出する手法を示している。文中の単語が同一概念を参照しているような文間にリンクがあるとしている。

Halliday と Hasan (Halliday and Hassan 1976) は、表層的な文間のつながりを表す指標として、5 種類の結束性 (cohesion)、すなわち、指示 (reference)、代入 (substitution; たとえば、‘a

new one' における 'one', 'do so' における 'do' などを用いた照応), 省略 (ellipsis), 接続 (conjunction), 語彙的結束性 (lexical cohesion) をあげている. 語彙的結束性は, 関連性のある語彙が用いられることで, 複数の文間の意味的なつながりが明示される場合であり, Skorokhod'ko が文間にリンクを与えたのはこの場合に相当すると考えられる.

Hoey(Hoey 1991) は, この語彙的結束性の情報を利用し, 文間で単語によるつながりが多いほど, 文間のつながりが強いと考え, 他の文とのつながりの強さに基づき, 要約を作成する手法を示している.

また, 語彙的結束性の情報を, 互いに関連のある単語のつながりである語彙的連鎖 (lexical chain)(Morris and Hirst 1991) として計算し, それを要約の知識源として用いる研究としては, 佐々木ら (佐々木, 増山, 内藤 1993)⁵, Barzilay と Elhadad(Barzilay and Elhadad 1997), 望月ら (望月, 岩山, 奥村 1998) がある.

奥西ら (奥西, 吉見, 山路, 福持 1998) は, テキストのタイトルを最も重要な文と考えた上で, 重要な文とのつながりが強い文を重要と考える重要文抽出手法を提案している. 文の重要度は, 先行文とのつながりの強さと, 先行文の重要度の積で計算されるが, 文間のつながりの強さは, 同一単語の出現により得られる語彙的なつながりの情報などを基に計算される.

Mani ら (Mani and Bloedorn 1997) は, テキスト中の単語などがノードであり, その間の隣接性, 構文的関係, 共参照関係, 語彙的類似性などの関係をアークで表現したグラフでテキストを表現し, このグラフ中での活性値の伝播により, 高い活性値を得た単語, 句, 文を重要とみなす重要文抽出手法を示している. 検索結果のテキストの適切性判定に要約を用いる評価方法で, 活性値が上位 5 文の要約 (提案する手法による) と原文を比較した結果, 精度を落とすことなく, 短い時間で適切性判定ができることを示している.

Mani ら (Mani, Bloedorn, and Gates 1998) は, 上で述べた, 単語間のつながり (結束性) に基づく要約手法と, 節, 文間の関係を解析したテキスト構造に基づいた要約手法を比較している. テキスト構造を利用した手法としては, (4) で紹介した Marcu の研究を利用している. 人間の選択した重要文との一致度で評価した結果, テキスト構造を利用した手法の方がわずかではあるが良い結果を得られることを示している.

(7) テキスト中の文間の類似性の利用

情報検索の分野では, テキスト (や, その断片) を, その中に出現する単語の重みのベクトルとして表現することが多い. このような表現を用いると, テキスト間の類似度は, テキストを表現するベクトル間の内積等で計算することができる⁶.

これと同様に, テキスト中の文 (段落) を一単位として, それらの間の類似度を計算し, こ

⁵ 佐々木らは, 語彙的連鎖ではなく, 結束チャートと呼んでいる.

⁶ ここで利用される単語の重み付け手法及び, ベクトル間の類似度計算手法の数々については (Salton 1989) を参照して欲しい.

の類似度を文(段落)間のつながりの度合と考え、この情報を基に、重要と考えられる文(段落)を抽出する手法がいくつか提案されている。これらの手法は、文(段落)間で共通の単語が出現する度合に基づき計算される文(段落)間のつながりにより、重要文(段落)抽出を行なっていると考えられる。

Saltonら(Salton, Singhal, Buckley, and Mitra 1996)は、段落をノードとし、(ある閾値以上)類似度の高い段落同士をリンクで結んだ text relationship maps をまず生成し、そこから重要段落を抽出する。Mitraら(Mitra, Singhal, and Buckley 1997)は、この手法を、人間の抽出した重要段落との比較により評価している。百科辞典の見出し語に対する項目を対象に、2人の人間の抽出した段落の一致の度合46%に対し、人間の段落とシステムの段落の一致の度合は、ほとんど変わらないとしている。先頭20%の段落を抽出した場合が提案手法を上回る精度を得ているが、これは、百科辞典が新聞同様、先頭に見出し語の定義など、主要な内容を含んでいるからと考えられる。

亀田(亀田 1996)は、2文間にどのくらい共通な単語(キーワード)が現れるか⁷に基づいて計算した文間関連度(の平均)と、ある文が他の文とどの程度広く関連があるかというカバレジに基づいて文の重要度を計算する手法を提案している。福本(福本 1997)も、文を単語の(重みの)ベクトルとして表現し、ベクトルの内積で文間の結合度を計算し、結合度の高い文を順に抽出する手法を提案している。

2.2 複数の表層の手がかりを統合して用いる要約手法

前節で述べたように、重要箇所抽出には、これまでさまざまな情報が用いられてきている。これらの情報のうち一つを用いた手法も提案されているが、一般に複数の情報を同時に用いた方が精度を改善できるとの考え方に基づき、複数の情報を統合して重要箇所抽出に用いる研究が数多く見られる。本節ではそのような研究を概観する。

山本らのシステム GREEN(山本, 増山, 内藤 1995)は、以下にあげる2つの情報を用いた要約システムである。

文の種類 人手で作成したパターンにより文の種類を分類し、著者の主張等を述べた文(見解文)のみを抽出する、

テキスト中での位置 テキスト、段落の先頭、最後の文を抽出する

この他に、重要文として抽出される文の先頭に指示語や接続詞が出現する場合、あるいは、抽出される文の主語が省略されている場合、前文との結束性が強く、単独で要約中に存在すると不自然であると考え、前文も要約に追加する処理を行なっている。また、重要文中で比較的重要度が低いと考えられる連体修飾句の削除を行ない、さらに要約を短くすることを試みている。要約の評価は、要約の自然さ、内容の適切さに関する被験者へのアンケートにより行なっている。

⁷ 厳密には、どの程度単語の部分文字列に重複があるか

亀田 (亀田 1997) は, 類似した単語 (キーワード) が 2 つの文に共に出現する度合に応じて計算される文間関連度に基づいて, テキスト中の文の重要度を計算する, 以前の手法 (2.1 節 (7) 参照)(亀田 1996) に, 段落や見出しの情報などを追加して重要文の抽出を行なう手法を提案している. この手法では, 文間関連度に基づく文自体の重要度の他に, 段落間関連度を用いて段落の重要度のランキングを行ない, 段落の重要度を文の重要度に加味する (重要度の高い段落中の文に, より大きい重要度を付与する). さらに, 見出し文と関連する文 (2.1 節 (3) 参照) や, 重要性を示す機能語句を含む文 (2.1 節 (5) 参照) を重要視する補正をこれに加え, 最終的な文の重要度を計算する. 被験者の抽出した重要文との一致の度合により評価を行ない, 以前の手法, 市販のソフトウェアとの比較の結果, 提案した手法が優れていることを示している.

このように, テキスト中の箇所 (文, 段落) に重要度を付与する情報を複数統合する際には, 情報の統合の方法が問題となる. これまでの研究では, 個々の情報により付与された重要度に, それぞれの情報の重みをかけたものを足し合わせ, 全体としての重要度とする手法がよく用いられている⁸.

このような重要度計算の手法を用いる場合, それぞれの情報に対する重み付けは, これまで人手でその重みを調整する手法が取られることが多かった.

Edmundson (Edmundson 1969) は, 2.1 節で上げた, いくつかの情報を人手で組み合わせた重要文抽出手法による実験を行ない, 複数の情報を組み合わせることで, より良い結果が得られることを示している. 個々の情報を用いた手法では, 「位置情報」, 「手がかり表現」, 「タイトル等の情報」, 「出現頻度」の順に精度が良い. しかし, 最初の 3 つの情報の組合せがもっとも良い結果を得られたと報告している.

間瀬ら (間瀬, 大西, 杉江 1989) は, 2.1 節で上げた情報 「タイトル等の情報」, 「出現頻度」, 「手がかり表現」, 「段落中での位置情報」⁹ に対応するパラメタおよび, 主題 (助詞 「は」 が後置する名詞), 文間の接続詞 などの情報を基にしたパラメタを組み合わせて利用する重要文抽出手法を示している. 文の重要度は, 各パラメタのスコアに, 人手で決定したパラメタの重みをかけたものを総和することで計算している. さらに, 抽出した重要文に指示語が存在する場合, 先行詞を含む文を推定し同時に抽出したり, 不要と考えられる接続詞や副詞を削除するなどの後処理を加えている.

これに対し, 近年要約文集を訓練コーパスとして, 機械学習手法などを用いることにより, 複数の情報の統合方法を最適化する研究が盛んになってきている. 複数の情報の重みを自動的に決定するのはその一例と言える. このような研究は, その学習法の違いから次の 3 つに大別できる.

8 ある文に対して複数の情報が矛盾する判断をする場合や, また, 複数の情報間に依存関係がある場合などを考慮すると, 単純に複数の情報をスコアとして統合するのが適切なのかという疑問はある.

9 間瀬らは, 段落を表層的な情報を基に分割した, セグメントという意味的なまとまりも考慮に入れている.

- 確率を用いたベイズ推定手法 (Kupiec, Pedersen, and Chen 1995; Jang and Myaeng 1997; Teufel and Moens 1997)
- 重回帰分析を用いた手法 (Watanabe 1996)
- 決定木学習を用いた手法 (野本, 松本 1997; Aone et al. 1997)

Kupiecら (Kupiec et al. 1995) は, 重要文抽出を統計的な分類問題とみなし, あらかじめ人手で選択した重要文を訓練集合とし, 文が重要文集合に含まれるかどうかの確率を与える分類関数を求めておき, 重要文はこの確率により文を順序付けることで抽出する重要文抽出手法を提案している. 具体的には, 属性 (重要文抽出のための情報) 集合が与えられた時の, 文 s が要約 S に属す確率 $P(s \in S | F_1, F_2, \dots, F_k)$ を計算する. 属性集合としては, 2.1 節で述べた「出現頻度」, 「手がかり表現」, 「位置情報」に対応するものの他に, 「短い文は重要文になり難い」, 「固有名詞は重要であることが多い」という仮定を導入するための属性も利用している. 人手で属性の重みの決定を行なった Edmundson の研究と基本的に結果が一致している. また, テキストの先頭の文を抽出する手法をベースラインとすると, ベースラインの精度 24% に比べ, 42% 平均の適合率を得ている.

Watanabe (Watanabe 1996) は, 属性集合として, 2.1 節で述べた「出現頻度」, 「位置情報」に対応するものの他に, 「時制情報」(現在/過去), 「文のタイプ」(事実/筆者の主張/推測), 「前文との接続関係」(理由/例示/逆説/並列/...) を利用している. これらの属性それぞれの各文に対するスコアに属性の重みをかけたものの総和を文の重要度とするが, 属性の重みを訓練データから重回帰分析を行ない求めている.

Nomoto (野本, 松本 1997), Aone ら (Aone et al. 1997) はそれぞれ, 人手による要約文を訓練データとし, 決定木学習アルゴリズム C4.5 (Quinlan 1993) を用いて学習した決定木により, 文を重要文/非重要文に分類し, 重要文抽出を行なう手法を提案している. Kupiec らの手法と同様に, テキスト中の文が重要文集合に属すかどうかを分類する分類器を学習するが, ここでは, 確率ではなく決定木を学習する. 決定木学習は, あらかじめ分類済みの訓練データに属性情報を付加しておき, そのデータを正しく分類できるようなルール集合を決定木の形で学習することになる.

Mani ら (Mani and Bloedorn 1998) は, 異なる学習手法を用いた要約の精度の比較を行なっている. 「位置情報」, 「単語の出現頻度」, 「タイトル等の情報」, 「文間の結束性の情報」を属性として用いている. 対象テキストとしては, Teufel らと同じ e-print archive 中の学術論文 198 編および著者の記述した要約を利用している. 要約の長さは平均で原文の 5% であった. 決定木学習を用いた結果から, 位置情報, 単語の出現頻度の情報が有効である一方, 結束性の情報が効いていないことが明らかになった. また, 学術論文で有効と考えられる「手がかり語の情報」を利用していないにもかかわらず, 比較的良好な結果を得ている. 結果から, 重回帰分析を用いた手法, 決定木学習を用いた手法は, ほぼ同等な性能を得ているとしている.

2.3 重要文抽出に基づく要約の問題点

これまでの要約研究として、テキスト中の重要箇所の抽出とその連結による生成に基づくものを説明してきた。

説明してきた手法のうち、単語の出現頻度を用いた手法や、単語間のつながりの情報を用いた手法では、テキストが複数の話題から構成されている場合、話題ごとに語彙の出現傾向が変わるため、十分な精度が得られない可能性がある。同様に、位置情報を用いた手法も、十分に機能しない可能性がある。このような場合は、テキストを何らかの手法(たとえば、(Hearst 1994; 望月, 本田, 奥村 1999))で話題ごとに分割し、話題ごとに重要文を抽出する必要がある(仲尾 1998)。

また、重要箇所抽出に基づく要約手法の問題点としては、1) 抽出した文中に代名詞などが含まれている場合、その先行詞が要約文中に存在する保証がないこと、2) テキスト中の色々な箇所から抽出したものを単に集めているため、抽出した複数の文間のつながり(首尾一貫性)が悪いことが指摘されている。

Paice(Paice 1990)は、テキスト中のキーとなる重要な文を抽出するために用いられる表層的な情報の概説(2節)の後、重要文抽出による要約作成の問題点として、上のような点を指摘し、その問題点の解決法について述べている。照応詞を含む文の前の数文を要約に追加したり、接続詞は削除したり、動詞の時制や態は調和がとれた流れにしたりすることで、部分的な解決が実現できることを示している(3節)。また、テキストの構造の把握が要約文作成に重要であることを4節で述べている。

2.2節で紹介した、山本らのシステム GREEN では、上に述べた問題点に対し、重要文として抽出される文の先頭に指示語や接続詞が出現する場合、あるいは、抽出される文の主語が省略されている場合、単独で要約中に存在すると不自然であると考え、前文も要約に追加することで対処している。

2.1節(2)で紹介した、Brandowらのシステム ANES では、文の先頭に照応詞が出現する場合、その文を重要文として選択しないことや、また、段落の先頭でない文(2, 3番目の文)が選択された場合、段落中のそれらの文の前の文を要約に追加するなどの方法で、要約の結束性を増すことを試みている。

最後に、2.1節で説明した、重要箇所抽出のための様々な情報がどれも、すべてのジャンルのテキストで有効に機能するというわけではないことに注意しておきたい。新聞記事、学術論文など、テキストのジャンルが異なれば、有効な情報は異なるのが当然である。これまであまり議論されてこなかったが、テキストのジャンルと、重要文抽出に有効な情報の関係について、今後詳細に検討する必要があると考えられる。また、情報の組み合わせ方に関しても、テキストのジャンルによって、最適化した組み合わせ方での精度がばらつくこと、最適化した際に利用される属性集合が異なることなどが報告されている(野本, 松本 1997; Aone, Okurowski, and

Gorlinsky 1998).

3 抽象化, 言い換えによる要約手法

これまで述べてきた研究はどれも, 要約文 = 抽出したテキスト中の重要箇所の連結」という考え方に基づいていた. これは, 要約を「原文から (何も変えずに) 抽出した extract(抜粋)」と見なしているとも言える. これに対し, 言い換えたり, 合成したりすることで, 原文の内容を表現し直し要約 (abstract) として生成する」試みが近年いくつか見られるようになってきた.

この abstract の生成のためには, 通常の extract(テキスト中の重要概念の抽出) 以外に, 抽出した概念の統合, 生成の過程が必要である. 概念の統合は, 抽出された複数の重要概念を, 何らかの知識を用いて, より高い階層の概念にまとめることである. これにより, テキスト中の重要概念は, より少ない数の概念で表されることになる. 概念の統合には, 概念階層やスクリプトといった知識が必要となる. Hovy らの SUMMARIST(Hovy and Lin 1997) システムは, WordNet を概念階層として利用し, このような概念統合を実現している. Hovy らは,

John bought some vegetables, fruit, bread, and milk.

のような文を, 概念階層を用いて,

John bought some groceries.

のように言い換える処理を, 概念階層を用いた概念統合の例として示している.

Kondo と Okumura(Kondo and Okumura 1997) は, Hovy らの用いている概念階層以外に, EDR 単語辞書中の単語の定義文をスクリプト知識と見なし, 定義文を利用して, 概念統合を実現する手法を示している. 定義文は見出し語の説明であるため, 逆に見出し語は定義文中に出現する動作系列の簡潔な言い換えになっていると考えられる. たとえば,

説得する: よく話して納得させる

納得する: 物事を理解して承認する

承認する: 相手の言い分を聞き入れる

のように, それぞれの単語の定義文が与えられている場合,

私は彼女に事情を話した.

彼女は私の言う事を理解し,

聞き入れてくれた.

のような文は, 上の 3 つの定義文を利用して, 定義文中の動作系列を再帰的に見出し語に言い換えることにより,

私は彼女を説得した.

のような文に要約できると考えられる.

4 ユーザに適応した動的な要約手法

これまでの要約研究は主に、対象となるテキストの情報を基に、要約は静的に決定できるという考え方で進められてきたように思われる。これに対して、近年、要約の利用される状況でユーザの要求に適合した要約を動的に作成する必要があるという考え方に基づいた研究が開始されている (Ochitani et al. 1997)。

たとえば、情報検索において、ユーザがクエリを入力し、検索されたテキストが適切かどうかを判断する際に要約を用いる場合を考えると、要約はユーザが入力したクエリに即したものになっている必要があり、これまでのように、テキストの内容のみから作成していたのでは必ずしも十分ではないと言える。また、ユーザの持つ予備知識の程度に応じて、出力する要約の詳細さ、長さは可変であるべきであると考えられる。

岩山ら (岩山, 徳永 1999) は、テキスト分類において、長いテキストを対象とする場合、テキスト中に複数の話題が含まれるなどにより、テキスト全体を単位とするよりも、テキスト中の断片を処理対象とした方が精度が改善できるという考えに基づき、パッセージ分類という手法を提案している。このパッセージ分類では、分類されるカテゴリと関連の強いパッセージをテキスト中から抽出しており、カテゴリを観点とした動的な要約作成を行なっていると言える。同様に、近年テキスト検索において注目されているパッセージ検索は、ユーザの入力したクエリに関連する、テキスト中のパッセージを抽出し、それを基に検索するわけなので、動的な要約作成を行なっていると言える (たとえば、(望月, 岩山, 奥村 1999))。

Tombros ら (Tombros and Sanderson 1998) は、「テキストのタイトル情報」、テキスト中での位置情報、テキスト中の単語の出現頻度」に基づいた、従来通りの文の重要度に、クエリ中の単語が文中に出現する頻度に応じたスコアを加味することで、クエリに依存した重要文抽出手法を実現している。また、この手法で抽出された query biased summary の、情報検索時における有用性を、検索されたテキストのリストから適切なテキストを同定するタスクにおける、被験者の速さ、精度を計ることで評価している。query biased summary と、テキストの先頭数文を抽出した要約を比較することで、query biased summary の有用性が示せたとしている。

塩見ら (塩見, 徳田, 青山, 柿ヶ原 1998) も、「テキスト中の単語の出現頻度」に基づいた従来通りの文の重要度に、クエリ中の単語が文中に出現する頻度に応じたスコアを加味することで、クエリに依存した重要文抽出手法を実現している。Tombros らと同様、情報検索時における有用性を、BMIR-J1 を利用して評価した結果、要約率 20% の要約文において、従来の単語の出現頻度に基づいた手法と比較することで、提案する手法の有効性を示せたとしている。

5 複数テキストを対象にした要約手法

これまで、単一テキストの要約作成に関する様々な手法について述べてきた。要約対象が複数テキストの場合、単一テキストの要約とは別に考慮すべき点が出てくる。まず、要約対象となる複数のテキストをどのように収集するのか。また、収集してきたテキスト間で内容が重複する場合、従来の単一テキスト要約の手法を個々のテキストに適用し並べただけでは、個々の要約の記述が重複する可能性があり、冗長で要約として適切ではない。そのため、冗長な箇所(テキスト間の共通箇所)をどのように検出し削除するかが問題となる。一方、冗長な箇所を削除しても複数テキストの要約文書としてはまだ十分であるとは言えない。複数のテキストを要約するとは、それらのテキストを比較し要点をまとめることであり、そのためにはテキスト間の共通点だけでなく相違点も明らかにすることが必要であると考えられる。さらに、要約文書を作成するためには、検出されたテキスト間の共通点や相違点を並べ、使用する単語の統一、接続詞の付与等の読み易さを上げるための処理を行う必要があると考えられる。従って、複数テキスト要約のポイントは次のようにまとめることができる。

- | | | | |
|---|------------------------------|---|----------------------|
| { | (a) 関連するテキストの自動収集 | { | (b) - 1 重要箇所の抽出 |
| | (b) 関連する複数テキストからの情報の抽出 | | (b) - 2 テキスト間の共通点の検出 |
| | (c) テキスト間の文体の違い等を考慮した要約文書の生成 | | (b) - 3 テキスト間の相違点の検出 |

複数テキスト要約に関するこれまでの研究には以下のものがある。Yamamotoら (Yamamoto, Masuyama, and Naito 1995), 稲垣ら (稲垣, 早川, 田中 1998), 柴田ら (柴田, 上田, 池田 1997), Radevら (Radev and McKeown 1998), Maniら (Mani and Bloedorn 1997) はいずれも、複数の新聞記事を対象に研究を行っている。また、難波ら (難波, 奥村 1999) は学術論文を対象に研究を行っている。

複数新聞記事を要約対象とした、これまでの研究は、次に示す大きく2つに分類される。

- (i) ある事件について書かれた記事とその続報記事から要約を作成する,
- (ii) ある事件に関する複数の情報源(新聞社)の記事を要約対象とし、要約を作成する。

柴田らは、Fitという検索システムに文章融合機能を埋め込み、自動分類された新聞記事の融合を試みている。柴田らは複数の情報源から得られる記事からの要約作成を試みている((ii))。柴田らは、ある事項に関連する記事が複数の情報源から得られた場合、記事間の共通箇所を抽出することが関連記事の重要箇所を抽出することであると考えている。手法としては、出現頻度の低い形態素が異なるテキストで出現する場合それを含む文は重複文(テキスト間の共通点)の可能性が高いと考え、重複文を同定し、その片側を用いて要約作成を行っている。稲垣らも、

柴田らと同様, ある事件について, 複数の情報源 (新聞社) から発行された記事から要約を作成する手法を提案している ((ii)). 記事間の共通箇所を抽出することが関連記事の重要箇所を抽出すること」という考え方は, 柴田らと同じであると言える.

Radevらは, 情報抽出手法により生成されたテンプレートを用いて, 複数の新聞記事の要約を試みている. Radevらも, 柴田ら, 稲垣らと同様に複数の情報源から得られる新聞記事を要約対象としているが, 同じ情報源から得られる続報記事についても考慮している ((i)(ii)). 要約対象はテロリストに関する記事にあらかじめ限定されている. まず, 情報抽出手法によりテンプレートに犯人, 犠牲者, 事件のタイプなどの計 25 の情報を抽出する. 次に, テンプレートを用いて要約を作成する. 一般に, 古い記事では不完全であった情報が続報記事中で明らかになった場合, 要約作成には新しい情報を優先させる必要がある. また, 同じイベントが異なる情報源でレポートされ, それらが互いに不完全な情報であるならば, 組み合わせることで, より完全な情報が得られる場合がある. このような点を考慮し, 複数記事から得られた情報の共通点, 相違点を考慮し統合するための 7 種類のオペレータを準備し, 要約作成を行っている.

Maniらは, 関連のある一組のニュース記事の要約を試みている. 要約対象となる一組の記事が (i) であるのか (ii) であるのかについては, 論文中では明らかにしていない. Maniらは, 2.1 節 (6) で紹介したように, 個々のテキスト (新聞記事) を関連する語句 (ノード) の間にリンクを張ったグラフで表現している. そして, 活性伝搬により, テキストの話題と関係するノード集合 (サブグラフ) を検出する. 記事間でそれらのサブグラフを照合することで, テキスト間の共通点と相違点の抽出を行っている.

難波らは, 特定分野の複数の論文からサーベイ論文を自動作成することを目指しており, その第 1 歩としてサーベイ論文作成支援システムを構築している. 難波らは, 論文間の参照情報に着目し, 参照情報を用いて論文間の共通点や相違点を明らかにする手法を提案している. 参照情報とは, 論文中で, 参照先論文について記述している箇所 (参照箇所) から得られる情報のことで, 参照先論文の重要点や, 参照元と参照先間の相違点を明示する有用な情報が得られる. 難波らは, 参照箇所を cue word を用いて解析し, 論文の参照・被参照関係にリンク属性 (参照タイプ) を付与している. 特定のリンク属性が付与された参照関係を辿ることで, ある特定分野の論文を自動的に収集するのに近い処理を実現している. こうして収集された論文集合の参照関係のグラフや, 個々の論文のアブストラクト, 参照箇所を示すことで, ユーザに関連論文の共通点や相違点を明示できるため, サーベイ論文作成に有用であると考えられる. しかし, これらの情報を用いてサーベイ論文を自動的に作成するには至っていない.

6 文中の重要箇所抽出, 不要箇所削除による要約手法

これまでの要約手法の多くは, テキスト中の重要な文あるいは段落を抽出することで実現されていた. しかし, 文単位の抽出では, 重要でないとして捨てられる情報の単位が文であるこ

とから、要約を作成する際に、情報が大きく欠落する可能性がある。そのため、文単位で抽出することでテキストを短くするのではなく、一文ごとに重要でない箇所を削り(あるいは、重要な箇所を抽出し)、情報をなるべく減らさずに、テキストを短く表現し直す要約手法が近年提案され始めている。これらの手法は、段落、文、節を単位とした重要箇所抽出ではなく、句、文字列を単位とした重要箇所抽出(不要箇所削除)とすることができる。

これらの手法のもう1つの特徴として、具体的な利用目的を想定した要約研究として手法が提案されていることが上げられる。その一つが、文字放送、字幕を作成することを想定した要約手法としての、文の短縮である。

文字放送、字幕を作成することを想定した場合、文字放送、字幕では体言止め、漢字熟語などを多用した、固有の表現が可能であること、また、文字放送、字幕用要約の場合、通常の要約と比べると、要約の長さをそれほど短くする必要がないことなどから、不要と考えられる文字列を削除したり、表現をより簡潔な別の表現に言い換えるなど、表層の文字列に関する処理で、ある程度文を短縮することが可能である。文末のサ変動詞を体言止めにする(「月中に解散します」→「月中に解散へ」)、文末の丁寧の助動詞を削除する(「余震が相次ぎました」→「余震が相次いだ」)などのような変換規則を用意し、文に対し変換規則を繰り返し適用することで、文はより短い文に変換される。

若尾ら(若尾, 江原, 白井 1997)は、実際のニュース番組中の字幕を用いて、人手で作成されている字幕とニュース原稿を比較することで、字幕用の要約手法の分析を行なっている。要約手法として、表層の文字列の情報のみで可能な手法のみを取り上げ、5つに分類している。また、各手法の使用頻度、削減される文字数なども調査している。

山崎ら(山崎, 三上, 増山, 中川 1998)も同様な手法を提案し、要約率91.2%を得たとしている。若尾ら, 山崎らがともに、元原稿と字幕を人手で分析し、要約のための規則を作成しているのに対し、加藤(加藤 1998)は、この要約のための知識を自動的に獲得する手法を提案している。原文となるニュース文原稿と、要約文となる文字放送原稿のペアからなるコーパスを利用して、要約知識を自動獲得している。ペアとなる文間の対応をDPマッチングにより単語単位でとり、その後対応のとれなかった差分の部分を要約のための変換規則として獲得している。

また、近年モバイルコミュニケーションが脚光を浴びているが、限られた通信・表示リソースしか持たないモバイル端末へのテキスト表示のための要約技術の研究も開始されている(稲垣, 早川, 井上, 田中 1998)。この場合も、重要文抽出ではなく、情報をなるべく欠落させず表示する必要があることから、字幕作成の場合と同様な技術が用いられる。

一方、テキストを構文解析し、その結果を利用して文中の重要箇所を抽出する手法がいくつか提案されているが、これらはいずれも、人間がテキストを走り読み(skimming)することを支援するために提案されている手法である。

亀田(亀田 1995)の日本語文書読解支援系QJRのskimming支援では、文書の速読を支援

する目的で、簡易日本語解析系 QJP による文の係り受け解析の結果を基に、文の骨格となる文節群のみを強調表示する機能を提供している。Grefenstette(Grefenstette 1998) は、視覚障害者が音声合成器を介してテキストを skimming するための文の単純化手法を提案している。Grefenstette が以前開発した shallow parser を基に文を構造化し、文の骨格となる、文中の重要箇所を抽出している。

これらの研究以前にも、山本ら (2.2 節参照)、Mahesh(7 節参照) などのように、抽出した重要文中の不要箇所を削除し、さらに要約文を短くすることを目的として、構文解析を利用した文中の重要箇所抽出は行なわれている。この場合も、構文解析結果から、連体修飾句、埋め込み文、従属節などを不要として削除している。

7 要約の表示方法について

これまでの要約研究においては、要約は、原文同様テキストとして、出力されることが一般的であったと言える。

しかし、2.3 節で述べたように、重要箇所抽出に基づく、伝統的な要約手法では、出力される要約が、テキストとしてのまとまりを十分構成しておらず、読み易さの点で問題があることが指摘されている。また、4 節で触れたように、要約の長さは、ユーザが自分の関心に応じて自由に変えられるようになっていくことが望ましいという指摘もある。

このような立場から、近年要約を、単なるテキストとしてではなく、他の形でユーザに表示する試みが行なわれ始めている。

Mahesh(Mahesh 1997) は、要約過程を、テキストからの hypertext の構成過程ととらえ、重要な箇所が前面にあり、そこからリンクをたどることで、より詳細な情報が段階的に得られるような枠組を提案している。従来の研究が(ある決まった要約率の)要約と、(要約する前の)テキスト全体という 2 つの要素しか提示しなかったのに対し、ユーザが自分の関心に応じてリンクをたどることで、さまざまな要約率の要約を段階的に参照可能である枠組である。テキストからの重要文抽出は従来の手法を用いているが、その後処理として、抽出した文を(部分的に)構文解析し、埋め込み文、従属節などを削除することで、さらに要約文を短くすることを試みている。

Saggion と Lapalme(Saggion and Lapalme 1998) は、ユーザに適応した要約の出力法として、indicative な要約をまず表示し、そこから、要約が対応するテキスト中の断片がたどれるようにしておき、ユーザは自分の関心に応じ、そのリンクをたどることで、より情報量の多い(informative な)要約を見ることができるよう枠組を示している。

これらの研究はどちらも、重要文抽出手法で得られた要約の、テキストとしての首尾一貫性の欠落の問題に対して、要約を表示する際、表示した要約が 1 つのまとまったテキストでは本来なく、したがって、首尾一貫性がない可能性があること(前後の文は無関係であるかもしれな

いこと)を明示してやることで、部分的な解決を試みていると言え、興味深い。

8 要約の評価方法について

これまでの単一テキストを対象とする要約研究の多くは、人間の被験者の作成した要約文と、システムの作成した要約文を比較し、システムの要約文の再現率、適合率を評価尺度とした評価を行っていた。

しかし、人間においても要約というタスクは必ずしも容易ではなく、人間の被験者による要約が必ずしも高い割合で一致するとは言えない。また、この評価法の前提とする「ただ一つ正しい要約が存在する」という仮定が不自然であるという批判が以前からあり、要約システムの評価方法は再検討される段階にあると言える。

これに対して、Miikeら (Miike et al. 1994) は、要約を利用して人間がタスクを行なう場合の、タスクの達成率が間接的に要約の評価となるという考え方にに基づき、評価を行なっている。具体的には、情報検索における検索テキストの適切性の判断をする際に要約を用いることで、要約を評価し、タスクに要する時間と、検索の再現率、適合率で評価を行なっている。

DARPA Tipster プロジェクト (Phase III) の評価 (Hand 1997) においても、同様に、上の仮定の不自然さから、タスクに基づく評価方法が採用されている。Tipster プロジェクトでは、テキストの分類、情報検索における検索テキストの適切性の判断それぞれに要約を利用し、被験者のタスクに要する時間 (要約しないテキスト全体を用いた場合とも比較する)、タスクの精度により要約を評価する。

一方、間瀬ら (間瀬, 大西, 杉江 1989) は、原文を読んだ後および、その要約だけを読んだ後、原文の内容を問うテストを被験者に行ない、テストの得点比で要約の評価を行なっている。テストの問題作成の困難さが問題点として残るが、原文を伴わない状況での利用を想定した要約の内容の十分性の評価としては興味深い手法である。

このように、要約を用いて人間の被験者が何らかのタスクを実行する際の精度等を問題にするのではなく、要約を利用して何らかのタスクを実行する応用プログラムの精度を示すことで、間接的に要約の評価を行なうという試みも見られる。

隅田ら (隅田, 飯田 1997) は、抽出した要約文のみを索引およびスコアづけの対象としたテキスト検索システムの評価を行ない、テキスト全体を索引等に用いた場合に比べ、精度の向上が実現できることを示すことで、抽出した要約文がテキストの大意の把握に成功していることを間接的に実証している。良い要約が得られれば、重要な概念や単語のみが索引語として利用されるので、検索の精度が改善されるはずであるという仮定にこの評価は基づいている。

このような、要約文の内容に関する評価とは別に、要約文の「文章としての読み易さ」を評価する評価方法も考えられる (Minel, Nugier, and Piat 1997)。

2.1 節 (2) で紹介した、Brandowら、Wasson は、人間の受容可能性判断に基づいて要約を

評価している。受容可能性は、人間が、原文と照らし合わせて、内容と読み易さに関して、受容可能/不可能の判定を要約に対して行ない求められる指標である。

要約は、本来このように、内容に関する評価と、読み易さに関する評価の、両方の次元で評価されるべきであると言え、今後もより良い要約の評価方法の模索は続けられるものと考えられる。

1 節で述べたように、要約は一般に、その利用目的に応じて、次の 2 つのタイプに分けられることが多い (Hand 1997)。

indicative: 原文の適切性を判断するなど、原文を参照する前の段階で用いる

informative: 原文の代わりとして用いる

Miike ら, Tipster プロジェクトの評価は、要約を indicative なものとして評価していると言える。一方、間瀬ら, 隅田らの評価は、informative なものとしての要約の評価を行っていることになる。

ここで、Tipster プロジェクトにおける評価方法について、もう少し詳しく触れておく¹⁰。Tipster プロジェクトの評価法は、上にも述べたように、タスクに基づくものであるが、そのタスクは、以下の 3 つからなる¹¹。

Task	Summary type
categorization	generic, indicative
ad hoc retrieval	query-based, indicative
question-and-answer	query-based, informative

‘query-based’ 要約は、4 節で述べた、ユーザの要求に特化した要約、‘generic’ な要約は、特化しない要約を意味する。最初の 2 つのタスクでは、10%の要約率での要約と、開発者が「最も良い」と考える要約 (長さは問わない) を基に評価を行なう。

3 つ目のタスクでは、質問に対する解答の正当率で要約を評価する。質問はテキストごとに変わるものではなく、query で示された topic ごとに 5 つ用意される。ある topic に関する質問の正解は、質問作成者自身が、(質問に対する正解を与えていると判断した) 原文の passage を選ぶことで決定される。評価は、この passage を要約がどの程度含むかで人間が判断する。評価の指標である Answer Recall は、correct, partially correct, missing の 3 段階で判断される。要約の評価方法としては、上述した、間瀬らの手法と同様なものと考えられる。

一方、5 節で述べた、複数テキストを対象とする要約研究や、6 節で述べた、文中の重要箇所抽出による要約研究の評価は、研究が始まったばかりでもあり、十分な議論がなされてきていないと言って良い。

¹⁰ Tipster の SUMMAC に関する、簡単な報告が (福本, 関根, 江里口 1998) にある。

¹¹ Tipster では、これらのタスクに基づく評価以外に、受容可能性による評価も合わせて行なっている。

5 節で述べたように、複数テキストを対象とする場合、冗長な重複箇所を検出し、削除することが必要となるため、「冗長箇所をどの程度正しく削除できているか」(船坂, 山本, 増山 1996), 「テキスト間の類似箇所と相違箇所をどの程度正しく抽出できているか」(Mani and Bloedorn 1997) という観点での評価が行なわれている。また、難波ら(難波, 奥村 1999) は、要約に必要な記述内容(参照箇所)をどの程度正しく抽出できているかを評価している。しかし、複数テキストから作成された要約文全体に関する評価はこれまでなされておらず、どのような点を評価すべきかということも明らかではない。今後、作成された要約全体の評価について検討していく必要があると考えられる。

9 おわりに

テキスト自動要約に関する、これまでの研究動向を概観してきた。

自然言語処理の分野では、近年頑健な解析手法の開発が進んでいるが、これらの手法を用いた要約研究が今後も数多く提案されるようになると思われる。2.1 節(4)で述べた「解析したテキスト構造を利用した」要約手法も、テキスト構造を解析する頑健な手法が開発されて初めて実現可能な手法であり、また、照応解析を利用した要約手法(Boguraev and Kennedy 1997)など、頑健な文脈処理を利用した要約手法が今後盛んに研究されることと思われる。頑健な文脈処理を利用した手法は、2.3 節で述べたような、伝統的な重要箇所抽出による要約の問題点の解決にも貢献できる可能性が高いと言える。この他にも、複合語を抽出しそれを利用する、また、固有名詞を抽出し、そのタイプ(人名, 場所, 会社名など)わけを利用するなどして、要約手法をこれまでの単語に基づく単純なものから、より詳細な情報に基づくものに拡張し精度向上を図る試み(Aone et al. 1997)も増えていくと思われる。また、6 節で紹介したような、文中の不要箇所を削除したり、重要箇所を抽出したりすることによる要約手法では、頑健な(部分)構文解析手法の利用が不可欠であると考えられる。

最後に、本稿以外の過去の解説および参考文献を紹介しておく。(Cole, Mariani, Uszkoreit, Zaenen, and Zue 1997)の7.4 節に Spark Jones の簡単な解説がある。(Paice 1990)も、対象が論文中心ではあるが、2.1 節で述べたように、これまでの手法の解説を含んでいる。Information Processing & Management の Vol.31, No.5(1995)は自動要約(Automatic Summarizing)の特集号である。本稿では述べなかった、要約の生成過程に関する研究として、3 編の論文が収録されている。(Shapiro 1992)にも Alterman の解説がある。この解説は、対象が物語中心であり、領域知識を用いた手法に関してのみが説明されている。(Endres-Niggenmeyer 1998)の5 章は、計算機による要約手法をまとめた章となっている。人間の要約過程に関しては、(佐久間 1989), (Endres-Niggenmeyer 1998)などに詳しい分析がある。また、テキスト自動要約に関する Web page を最近作成した(<http://galaga.jaist.ac.jp:8000/pub/research/summarization/>)。興味のある方は参照して頂きたい。

謝辞

本稿を執筆する機会を与えて下さった, 本特集号編集委員の皆様にもまず感謝します. 本稿をまとめるに当たっては, 自然言語処理学講座に在籍する, 望月源君, 近藤恵子さん, 徳田昌晃君の協力が大きな助けとなりました. ここに記し, 感謝します. また, 本稿の予稿にコメントを寄せて下さった通信・放送機構 (TAO) の福島孝博氏, 日立中央研究所の小林義行氏に感謝します. Tipster の SUMMAC に関連する貴重な資料は, ニューヨーク大学の関根聡氏, ジャストシステムの野村直之氏に提供して頂きました. ここに感謝します.

参考文献

- Aone, C., Okurowski, M., and Gorlinsky, J. (1998). "Trainable, Scalable Summarization Using Robust NLP and Machine Learning." In *Proc. of the 17th International Conference on Computational Linguistics and 36th Annual Meeting of the Association for Computational Linguistics*, pp. 62–66.
- Aone, C., Okurowski, M., Gorlinsky, J., and Larsen, B. (1997). "A Scalable Summarization System Using Robust NLP." In *Proc. of the ACL Workshop on Intelligent Scalable Text Summarization*, pp. 66–73.
- Barzilay, R. and Elhadad, M. (1997). "Using lexical chains for text summarization." In *Proc. of the ACL Workshop on Intelligent Scalable Text Summarization*, pp. 10–17.
- Boguraev, B. and Kennedy, C. (1997). "Salience-based Content Characterisation of Text Documents." In *Proc. of the ACL Workshop on Intelligent Scalable Text Summarization*, pp. 2–9.
- Brandow, R., Mitze, K., and Rau, L. (1995). "Automatic Condensation of Electronic Publications by Sentence Selection." *Information Processing and Management*, **31** (5), 675–685.
- Cole, R., Mariani, J., Uszkoreit, H., Zaenen, A., and Zue, V. (Eds.) (1997). *Survey of the State of the Art in Human Language Technology*. Cambridge University Press. also available at <http://www.cse.ogi.edu/CSLU/HLTsurvey/>.
- Cowie, J. and Lehnert, W. (1996). "Information Extraction." *Communications of the ACM*, **39** (1), 80–91.
- Edmundson, H. (1969). "New methods in automatic abstracting." *Journal of ACM*, **16** (2), 264–285.
- Endres-Niggenmeyer, B. (1998). *Summarizing Information*. Springer-Verlag.
- 福本淳一, 安原宏 (1991). "日本語文章の構造化解析." 情報処理学会自然言語処理研究会報告. 85-11.

- 福本淳一 (1997). “文の結合度に基づく内容抽出手法.” 言語処理学会第3回年次大会発表論文集, pp. 321-324.
- 福本淳一, 関根聡, 江里口善生 (1998). “MUC-7, Tipster 参加報告.” 情報処理学会自然言語処理研究会報告, pp. 101-108. 127-14.
- 船坂貴浩, 山本和英, 増山繁 (1996). “冗長度削減による関連新聞記事の要約.” 情報処理学会自然言語処理研究会報告, pp. 39-46. 114-7.
- Grefenstette, G. (1998). “Producing Intelligent Telegraphic Text Reduction to provide an Audio Scanning Service for the Blind.” In *Intelligent Text Summarization*, pp. 102-108. AAAI Press. Technical Report SS-98-06.
- Halliday, H. A. K. and Hassan, R. (1976). *Cohesion in English*. Longman.
- Hand, T. (1997). “A Proposal for Task-based Evaluation of Text Summarization Systems.” In *Proc. of the ACL Workshop on Intelligent Scalable Text Summarization*, pp. 31-38.
- Hearst, M. (1994). “Multi-Paragraph Segmentation of Expository Texts.” In *Proc. of the 32nd Annual Meeting of the Association for Computational Linguistics*, pp. 9-16.
- Hoey, M. (1991). *Patterns of lexis in text*. Oxford University Press.
- Hovy, E. and Lin, C. (1997). “Automated Text Summarization in SUMMARIST.” In *Proc. of the ACL Workshop on Intelligent Scalable Text Summarization*, pp. 18-24.
- 稲垣博人, 早川和宏, 井上孝史, 田中一男 (1998). “モバイル端末の表示特性に応じたメッセージ要約方式の提案.” 情報処理学会第56回全国大会講演論文集 (分冊2), pp. 255-256. 5Q-2.
- 稲垣博人, 早川和宏, 田中一男 (1998). “類似意味内容の統合による伝達型電子化文書要約方式の提案.” 情報処理学会第57回全国大会講演論文集 (分冊2), pp. 253-254. 4R-11.
- 岩山 真, 徳永健伸 (1999). “確率モデルに基づくパッセージ分類とその応用.” 自然言語処理, 6 (3), 181-198.
- Jang, D. and Myaeng, S. (1997). “Development of a Document Summarization System for Effective Information Services.” In *Proc. of the RIAO97*, pp. 101-111.
- 亀田雅之 (1995). “日本語文書読解支援系 QJR の検討.” 情報処理学会自然言語処理研究会報告, pp. 57-64. 110-9.
- 亀田雅之 (1996). “疑似キーワード相関法による重要キーワードと重要文の抽出.” 言語処理学会第2回年次大会発表論文集, pp. 97-100.
- 亀田雅之 (1997). “段落間及び文間関連度を利用した段落シフト法に基づく重要文抽出.” 情報処理学会自然言語処理研究会報告, pp. 119-126. 121-17.
- 加藤直人 (1998). “ニュース文要約のための局所的な要約知識獲得とその評価.” 情報処理学会自然言語処理研究会報告, pp. 69-76. 126-10.
- Kondo, K. and Okumura, M. (1997). “Summarization with Dictionary-based Paraphrasing.”

- In *Proc. of the Natural Language Processing Pacific Rim Symposium'97*, pp. 649–652.
- Kupiec, J., Pedersen, J., and Chen, F. (1995). “A Trainable Document Summarizer.” In *Proc. of the 18th Annual International ACM-SIGIR Conference on Research and Development in Information Retrieval*, pp. 68–73.
- Kurohashi, S. and Nagao, M. (1994). “Automatic Detection of Discourse Structure by Checking Surface Information in Sentences.” In *Proc. of the 15th International Conference on Computational Linguistics*, pp. 1123–1127.
- Luhn, H. (1958). “The automatic creation of literature abstracts.” *IBM Journal of Research and Development*, **2** (2), 159–165.
- Mahesh, K. (1997). “Hypertext Summary Extraction for Fast Document Browsing.” In *Proc. of the AAAI Spring Symposium on Natural Language Processing for the World Wide Web*.
- Mani, I. and Bloedorn, E. (1997). “Multi-document Summarization by Graph Search and Matching.” In *Proc. of the 14th National Conference on Artificial Intelligence*, pp. 622–628.
- Mani, I. and Bloedorn, E. (1998). “Machine Learning of Generic and User-Focused Summarization.” In *Proc. of the 15th National Conference on Artificial Intelligence*, pp. 821–826.
- Mani, I., Bloedorn, E., and Gates, B. (1998). “Using Cohesion and Coherence Models for Text Summarization.” In *Intelligent Text Summarization*, pp. 60–67. AAAI Press. Technical Report SS-98-06.
- Marcu, D. (1997). “From Discourse Structures to Text Summaries.” In *Proc. of the ACL Workshop on Intelligent Scalable Text Summarization*, pp. 82–88.
- 間瀬久雄, 大西昇, 杉江昇 (1989). “説明文の抄録作成について.” 電子情報通信学会言語理解とコミュニケーション研究会報告, pp. 5–12. NLC89-40.
- Miike, S., Itoh, E., Ono, K., and Sumita, K. (1994). “A Full-Text Retrieval System with a Dynamic Abstract Generation Function.” In *Proc. of the 17th Annual International ACM-SIGIR Conference on Research and Development in Information Retrieval*, pp. 152–161.
- Minel, J., Nugier, S., and Piat, G. (1997). “How to Appreciate the Quality of Automatic Text Summarization? Examples of FAN and MLUCE Protocols and their Results on SERAPHIN.” In *Proc. of the ACL Workshop on Intelligent Scalable Text Summarization*, pp. 25–30.
- Mitra, M., Singhal, A., and Buckley, C. (1997). “Automatic Text Summarization by Paragraph

- Extraction.” In *Proc. of the ACL Workshop on Intelligent Scalable Text Summarization*, pp. 39–46.
- 望月 源, 岩山 真, 奥村 学 (1998). “抄録を利用したテキスト検索.” 言語処理学会第4回年次大会併設ワークショップ「テキスト要約の現状と将来」論文集, pp. 22–29.
- 望月 源, 岩山 真, 奥村 学 (1999). “語彙的連鎖に基づくパッセージ検索.” 自然言語処理, **6** (3), 101–126.
- 望月 源, 本田 岳夫, 奥村 学 (1999). “複数の表層的手がかりを統合したテキストセグメンテーション.” 自然言語処理, **6** (3), 43–58.
- Morris, J. and Hirst, G. (1991). “Lexical Cohesion Computed by Thesaural Relations as an Indicator of the Structure of Text.” *Computational Linguistics*, **17** (1), 21–48.
- 仲尾由雄 (1997). “見出しを利用した新聞・レポートからのダイジェスト情報の抽出.” 情報処理学会自然言語処理研究会報告, pp. 121–128. 117–17.
- 仲尾由雄 (1998). “文書の意味的階層構造の自動認定に基づく要約作成.” 言語処理学会第4回年次大会併設ワークショップ「テキスト要約の現状と将来」論文集, pp. 72–79.
- 難波英嗣, 奥村学 (1999). “論文間の参照情報を考慮したサーベイ論文作成支援システムの開発.” 自然言語処理, **6** (5), pp. 43–62
- 野本忠司, 松本裕治 (1997). “人間の重要度判定に基づいた自動要約の試み.” 情報処理学会自然言語処理研究会報告, pp. 71–76. 120–11.
- Ochitani, R., Nakao, Y., and Nishino, F. (1997). “Goal-Directed Approach for Text Summarization.” In *Proc. of the ACL Workshop on Intelligent Scalable Text Summarization*, pp. 47–50.
- 奥西稔幸, 吉見毅彦, 山路孝浩, 福持陽士 (1998). “ウェブ英文ページの速読支援.” 言語処理学会第4回年次大会発表論文集, pp. 572–575.
- Paice, C. (1990). “Constructing Literature Abstracts by Computer: Techniques and Prospects.” *Information Processing and Management*, **26** (1), 171–186.
- Quinlan, J. (1993). *C4.5: Programs for Machine Learning*. Morgan Kaufman Publishers.
- Radev, D. and McKeown, K. (1998). “Generating Natural Language Summaries from Multiple On-Line Sources.” *Computational Linguistics*, **24** (3), 469–500.
- Saggion, H. and Lapalme, G. (1998). “The generation of abstracts by selective analysis.” In *Intelligent Text Summarization*, pp. 130–132. AAAI Press. Technical Report SS-98-06.
- 佐久間まゆみ 編 (1989). 文章構造と要約文の諸相. くろしお出版.
- Salton, G. (1989). *Automatic Text Processing*. Addison-Wesley.
- Salton, G., Singhal, A., Buckley, C., and Mitra, M. (1996). “Automatic Text Decomposition Using Text Segments and Text Themes.” In *Proc. of the 7th ACM Conference on*

- Hypertext*, pp. 53–65.
- 佐々木一朗, 増山繁, 内藤昭三 (1993). “語彙的結束性に着目した文章抄録法の提案.” 情報処理学会自然言語処理研究会報告, pp. 65–72. 98-9.
- 関根聡 (1999). “テキストからの情報抽出–文書から特定の情報を抜き出す–.” 情報処理, **40** (4), 370–373.
- Shapiro, S. (Ed.) (1992). *Encyclopedia of Artificial Intelligence* (2nd edition). John Wiley and Sons, Inc.
- 柴田昇吾, 上田隆也, 池田裕治 (1997). “複数文章の融合.” 情報処理学会自然言語処理研究会報告, pp. 77–82. 120-12.
- 塩見隆一, 徳田克己, 青山昇一, 柿ヶ原康二 (1998). “視点を考慮した文書要約手法の提案.” 情報処理学会第 56 回全国大会講演論文集 (分冊 3), pp. 104–105. 3Y-8.
- Skorokhod'ko, E. (1972). “Adaptive method of automatic abstracting and indexing.” In *Information Processing 71*, pp. 1179–1182. North Holland.
- Sparck Jones, K. (1999, to appear). “Automatic summarizing: factors and directions.” In Mani, I. and Maybury, M. (Eds.), *Advances in automatic text summarization*. MIT Press. <http://xxx.lanl.gov/ps/cmp-lg/9805011>.
- 隅田英一郎, 飯田仁 (1997). “統計的な抄録法を使った情報検索.” 言語処理学会第 3 回年次大会発表論文集, pp. 353–356.
- Sumita, K., Ono, K., Chino, T., Ukita, T., and Amano, S. (1992). “A Discourse Structure Analyzer for Japanese Text.” In *Proc. of International Conference of Fifth Generation Computer Systems*, pp. 1133–1140.
- 鈴木康広, 栃内香次 (1988). “キーワード密接方式自動抄録法の改良.” 情報処理学会論文誌, **29** (3), 325–328.
- Teufel, S. and Moens, M. (1997). “Sentence Extraction as a Classification Task.” In *Proc. of the ACL Workshop on Intelligent Scalable Text Summarization*, pp. 58–65.
- Tombros, A. and Sanderson, M. (1998). “Advantages of Query Biased Summaries in Information Retrieval.” In *Proc. of the 21st Annual International ACM-SIGIR Conference on Research and Development in Information Retrieval*, pp. 2–10.
- 若尾孝博 (1996). “英語テキストからの情報抽出.” 情報処理学会自然言語処理研究会報告, pp. 77–83. 114-12.
- 若尾孝博, 江原暉将, 白井克彦 (1997). “テレビニュース番組の字幕に見られる要約の手法.” 情報処理学会自然言語処理研究会報告, pp. 83–89. 122-13.
- 若尾孝博, 江原暉将, 村木一至, 白井克彦 (1997). “テレビニュース番組電子化原稿を題材とした自動要約手法の大規模評価.” 情報処理学会自然言語処理研究会報告, pp. 31–36. 119-6.

- Wasson, M. (1998). "Using Leading Text for News Summaries: Evaluation Results and Implications for Commercial Summarization Applications." In *Proc. of the 17th International Conference on Computational Linguistics and 36th Annual Meeting of the Association for Computational Linguistics*, pp. 1364–1368.
- Watanabe, H. (1996). "A Method for Abstracting Newspaper Articles by Using Surface Clues." In *Proc. of the 16th International Conference on Computational Linguistics*, pp. 974–979.
- 山本和英, 増山繁, 内藤昭三 (1995). "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, 2 (1), 39–55.
- Yamamoto, K., Masuyama, S., and Naito, S. (1995). "An Empirical Study on Summarizing Multiple Texts of Japanese Newspaper Article." In *Proc. of the Natural Language Processing Pacific Rim Symposium '95*, pp. 461–466.
- 山崎邦子, 三上真, 増山繁, 中川聖一 (1998). "聴覚障害者用字幕生成のための言い替えによるニュース文要約." 言語処理学会第4回年次大会発表論文集, pp. 646–649.
- Zechner, K. (1996). "Fast Generation of Abstracts from General Domain Text Corpora by Extracting Relevant Sentences." In *Proc. of the 16th International Conference on Computational Linguistics*, pp. 986–989.

略歴

奥村 学: 1962年生. 1984年東京工業大学工学部情報工学科卒業. 1989年同大学院博士課程修了. 同年, 東京工業大学工学部情報工学科助手. 1992年北陸先端科学技術大学院大学情報科学研究科助教授, 現在に至る. 工学博士. 自然言語処理, 知的情報提示技術, 語学学習支援, 語彙知識獲得に関する研究に従事. 情報処理学会, 人工知能学会, AAAI, ACL, 認知科学会, 計量国語学会各会員. e-mail: oku@jaist.ac.jp .

難波 英嗣: 1972年生. 1996年東京理科大学理工学部電気工学科卒業. 1998年北陸先端科学技術大学院大学情報科学研究科博士前期課程修了. 同年同大学院博士後期課程, 現在に至る. 自然言語処理, 特にテキスト自動要約に関する研究に従事. 情報処理学会, 人工知能学会 各学生会員.