

文融合法に基づいた 放送ニュースリード文の具体化

田中英輝[†] 美野秀弥[†] 宮崎勝[†] 小早川健[†]
熊野正[†] 後藤淳[†] 加藤直人[†]

概要

本稿では放送ニュースのリード文を具体化する問題を取り上げる。これは、人手による要約が、リード文を他の文で具体化する手法を取っていることによる。これに対して、英語やドイツ語を対象に提案されている文融合法 (sentence fusion) に基づいたアルゴリズムを提案する。これはリードとその他の文を係り受け解析して、これらの間にある同一文節に係る句を置換、挿入するものである。この中では経験的に見いだした規則で置換、挿入を制御している。実験によって、このアルゴリズムを評価した結果、文法的には課題が残るものの、意味的には具体性が増すなどの効果が得られることがわかった。

Lead sentence Elaboration of Broadcast News with Sentence Fusion

Hideki Tanaka[†] Hideya Mino[†] Masaru Miyazaki[†]
Takeshi Kobayakawa[†] Tadashi Kumano[†] Jun Goto[†] and
Naoto Kato[†]

Abstract

We address the problem of revising the lead sentence in a broadcast news text to increase the amount of background information in the lead. We chose this problem as the professional abstractors of broadcast news text employ the strategy mentioned above. We propose an elaboration algorithm based on the sentence fusion method that has been applied to English and German so far. This method takes a pair of lead and another sentence that are parsed as input then finds the identical *bunsetsu* chunks among them. It next finds the modifier phrases to the same chunks and tries to elaborate the lead by inserting and substituting the modifier phrases in another sentence. We controlled the insertion and substitution operation with a set of heuristically obtained rules. Our experiments showed that most of the elaborated lead sentences were semantically correct though they contained some grammatical glitches.

1. はじめに

本稿では放送ニュースのリード文の具体化の問題を取り上げる。これは、放送ニュースの人手の要約が、リード文に欠如している背景などの具体的な情報を追加することで作られていることによる。

文書の要約は、重要文抽出法によって始まり[1]、今なお、それは基礎技術として使われている。しかし、この手法はよく知られているように、抽出された文同士に結束性が欠如してしまう問題、例えば、先行詞のない指示語の問題、を抱えている。

この結束性の問題を解決する手法として、draft and revision (原案と編集[a]) すなわち、抽出された文を編集する手法が提案された[2]。結束性の問題は特に複数文書要約で大きな問題となる。これについても多くの研究が行われており Otterbacher ら[3]は人手で修正された抽出結果を観察して結束性の問題を整理している。また Nenkova[4]は抽出された文の質的な向上を目指して名詞の参照関係を利用した編集手法を提案している。複数文書要約では類似した文書群を対象にすることが多く、結束性の問題以外にも、情報の重なりが問題となる。この問題に対して Barzilay and McKeown[5]は文融合法 (sentence fusion) と呼ばれる手法を英語に対して提案している。この手法はまず、対象となる文集合の依存構造を解析し、次に要約の基本 (basis) となる文を決める。この基本文に対して他の文の要素を情報が重ならないように付加して、最終的に不要句などを枝刈りする。この手法は Phillipova and Strube[6]によって改良されドイツ語に適用されている。

著者らの問題は、ニュース記事のリードの編集、すなわち単一日本語文書の draft and revision と考えられ、これに上記の文融合法を適用したものと見ることが出来る。実際、著者らの係り受け構造の対応付けは Phillipova and Strube[6]と同様であり、リード文が Barzilay and McKeown[5]の基本文と見ることが出来る。ただし上記の2手法が完全な要約作成を問題としているのに対して、著者らは、編集候補を見つける問題にこれを使っている。以下本論文の構成を示す。まず、2節で放送ニュースの人手の要約と本稿の問題を紹介する。3節では人手によるリードの修正実験とアルゴリズムの概要を示し、4節でその詳細を示す。続く5節では評価実験を報告する。

2. 放送ニュースの要約と要約システムの課題

著者ら[7]は、放送ニュース記事は3つの部分、リード、本記、追記からなることを報告している。これには、まず最重要な情報が冒頭の1ないし2文程度でリードとして簡潔に記される。リードには、必ずしも固有名詞や詳細な情報が記述されておらず、

[†] NHK 放送技術研究所
Science and Technology Research Laboratories of NHK

a) 本稿では、revision を表現の置換と挿入に限り、合わせて「編集」と呼ぶため、revision を「編集」とした。

会社名の代わりに「大手保険会社」などと記述されることがある。次にリードの詳細が本記と呼ばれる部分に3文ないし4文で記される。先のあいまいな会社名の実名はここに現れる。また、リードの内容が詳述される。最後に、その他の情報が追記と呼ばれる部分に1ないし2文で記される。なお、追記はないこともある。

著者ら[7]はまた、デジタル放送のための人手による要約手法を報告している。これによると、要約者は最初に記事の中のリード文を認定し、これを要約の基本とする。次に、リードの平均長が95文字程度で、求められる要約長が110文字程度[b]であることから、他の文（主に本記）の表現をリードにコピーアンドペーストすることで、リードをより具体化する。

以上の要約作成法はリードを draft とした、draft and revision の戦略に沿ったものと見ることができる。そこで、著者らは、放送ニュースの要約システムあるいは要約支援システムはこの戦略に基づくのが自然であると考えている。

この方針に従うならば、1) リード、本記、追記の認定、2) リードを具体化する編集候補の発見 3) 候補を使った最終的な要約生成、の3段階が必要になる。著者らはすでに1)の問題については決定木を使うことで92%程度の精度を確保できることを示している[8]。そこで、本稿では2)の問題を報告する。問題3)は今後の課題である。

3. 編集候補発見手法の概要

3.1 人手による表現編集実験

リードの編集候補の発見アルゴリズムを検討するため、2004年1月19日と20日から選んだ15記事を対象に、人手によってリードを他の文の表現で具体化する実験を行った。ここで許したのは、コピーアンドペーストによる表現の挿入と置換である。本稿ではリードの編集候補を探す相手の文集をソースと呼び（文はソース文）、挿入と置換を合わせて編集と呼ぶ。編集に使う表現は、連続した文節群に限った。ただし、連続文節の最終文節については、助詞の除去を許可した。

表1に得られた編集結果の一部を示す。リード列（第1列）の表現をソース列（第3列）の表現で編集したことを示す。リード列が空の場合は挿入で、そうでなければ置換である。また、それぞれの表現の係り先文節を示している（第2列と第4列）。挿入の場合の係り先とは、表現をリードに挿入した場合の最も適切な係り先文節である。この実験の結果、34件の編集が行われた。このうち助詞を取り除いた編集は5件であった（例を表1の第3行に示す）。以上より大半は表現そのままの編集であったことがわかる。次に、リードとソースの編集対象となった表現の係り先文節に着目し、その一致程度と、挿入、置換の数をまとめた結果を表2に示す。この中の自立語一致とは、表1の「会談し」と「会談しました」のような例である。

b) 画面サイズにより150文字程度のこともある。

表1 人手編集の例

リード	係り先	ソース文	係り先
	感染の	コイの大量死の原因となった「コイヘルペス」の	感染の
一部の旅客機で	見つかった	旅客機、MD811型機	一部の
ウルズガン州で	起きて	ウルズガン州のチャー・チノという村で	起きたもので
	会談し	中国国営の新華社通信によりますと	会談しました

表2 リードとソース文の係り先文節の一致と編集

	挿入	置換	計
完全一致	9	6	15
自立語一致	6	6	12
異文節	1	6	7
計	16	18	34

また、この表の「異文節」とは「見つかった」と「一部の」のような例である。異文節の先の係り文節を見ると同一の文節や、意味が類似した文節に係ることがわかった。例えば「一部の」という文節は「一部の→機体に→見つかった」と最終的には同一の「見つかった」に係っていた[c]。

また、「到着しました」と「入りました」（「サマーワに入りました」）のような同義語が係り先になる場合もあった。以上の観察から編集に使うことのできる表現は何らかの意味で“同一”の係り先を持つと仮定した。この点は4.1節で詳述する。

今回の調査では、置換は比較的容易に見つけられるが挿入を見つけるのは難しいことがわかった。置換表現を見つけるには、リードとソースにある類似表現を見つければよい。これに対して、挿入表現は、リードにない表現を探さなければならない。挿入の方が表現の発見に使える手がかりが少ない問題がある。

3.2 編集候補の発見方針

アルゴリズムを作るにあたっての課題は、編集候補の網羅的発見である[d]。これに対して、「編集に使える表現は“同一”の係り先を持つ」と仮定したことから、リードとソース文にあるすべての“同一”文節を探し、そこに係る表現から編集候補を探す方針を立てた。この手法は、置換、挿入の表現を探すのに利用でき、特に、手がかりの少ない、挿入表現の発見に使える利点がある。

c) 作業時には発見できなかったが「見つかった」に係る長い句を置換することも可能である。ここの編集にはあいまい性がある。

d) 著者らは2節で述べた段階3で、要約長に合わせて編集候補を取捨選択することを考えており、発見段階では意味のある候補を網羅したいと考えている。

3.3 編集候補発見アルゴリズム

本節ではアルゴリズムを示す。まず、本稿で使う述語を定義する。

3.3.1 トリガー文節

前節で述べた「同一文節」をトリガー文節と呼ぶ。この「同一文節」は編集候補発見の契機となることからこのように呼ぶ。具体的な同一性判定法は4.1節に示す。

3.3.2 最大句

トリガー文節に直接係っている文節が支配する最大の係り受け構造を最大句と呼ぶ。本アルゴリズムはこの最大句を編集候補の単位とする。なお、本稿では文節の連続を一般的に句と呼ぶ。

次に動作概要を図1とともに説明する。

3.3.3 動作概要 (Algorithm1, 表3)

入力係り受け解析されたリードとソース文のペアである。それぞれ L, S と記述する。(ステップ1)リードとソースの「同一文節」をトリガーペアとし T に格納する。(ステップ2,3) T の各トリガーペアに対して、トリガー文節の最大句を認定する。図1はトリガー文節「到着しました」を処理している例で、それぞれ2つの最大句が認定されている。

(ステップ4)類似した最大句同士を対応付ける。例えば、図1では「IAEAのチームが」と「IAEAの査察官5人が」が対応付けられ、それ以外は対応付けられていない。(ステップ5,表3)対応付け結果に応じて次の処理を行う。

表3-1 挿入処理: ソースだけにある句はリードへ情報を追加できるので、挿入候補として処理する。

表3-2 置換処理: 対応している句同士は類似しており、置換候補として処理する(置換せずに、挿入すると内容が重複することに注意)。

表3-3,4 処理なし: その他の場合はなにもしない。

以上から、ソース文の最大句は、対応付けのあと、挿入候補か置換候補となる。図1の場合で表3に従って置換、挿入を実行すると、「4日、」をリードに挿入して、「4日、 IAEAのチームが韓国に到着しました」という候補と、「IAEAのチームが」と「IAEAの査察官5人が」を置換して「IAEAの査察官5人が (IAEAのチームが) 韓国に到着しました」の2つの候補が得られる。なお、ここでの置換処理は句の分割処理を省略しており、詳細は4.4.1節を参照のこと。

4. アルゴリズムの詳細

本節ではアルゴリズム1のトリガー文節認定処理、句対応処理、挿入処理、置換処理を詳述する。まず、この後使う2つの基本的な評価指標を定義する。

Algorithm 1[e] (左の番号はステップ番号である.)

- 1: find all trigger pairs between L and S , and store them in T .
 $T = \{(l, s) ; l \approx s, l \in L \text{ and } s \in S\}$
- 2: for all $(l, s) \in T$ do
 find l 's max phrases and store in P_l .
 $P_l = \{p_l ; p_l \in \text{max phrase of } l\}$
- 3: do the same for trigger s
 $P_s = \{p_s ; p_s \in \text{max phrase of } s\}$
- 4: align phrases in P_l and P_s , and store result in A
 $A = \{(p_l, p_s) ; p_l \leftrightarrow p_s, p_l \in P_l, p_s \in P_s\}$
- 5: for all $(p_l, p_s) \in A$ do
 follow Table 3
 end for
- 6: end for

表3 ステップ5の動作

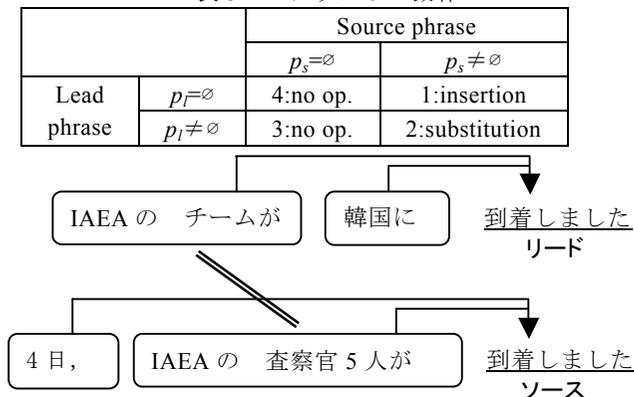


図1 編集候補発見アルゴリズムの概念図

文節類似度 t, s : 文節 \times 文節 $\rightarrow [0, 1]$

文節類似度 t は、文節を自立形態素の集合とみなして計算した Dice 係数とする。付属語を含めた文節類似度は s と記す。

句の含有率 D : 句 \times 句 $\rightarrow [0, 1]$

e) 記号 $a \approx b$ は文節 a と b がトリガーであることを示す。記号 $p \leftrightarrow q$ はフレーズ p と q が対応していることを示す。

句 X, Y に対して, X がどれくらい Y に含有されるかの量である. 文節類似度を用いて以下のように計算する. 絶対値記号は自立形態素の数を示す.

$$D(X, Y) := \frac{1}{|X|} \sum_{x \in X} \max_{y \in Y} (t(x, y))$$

4.1 トリガー文節認定処理

どのような文節がトリガーになるだろうか. これを定義する適切な術語が見あたらないが, 実験を通じた経験から, 同一の事実, 事物, 動作, 状態を参照する文節同士がトリガーの近似的な定義だと考える. このような照応関係は名詞を中心に議論されるが, 動詞その他の品詞からなる文節も本稿のトリガーに含まれる. 通常の参照あるいは照応の認定も困難な処理であることから, トリガーの認定はさらに困難だと考えた. そこで, 今回は表層が同一の文節をトリガーの基本として, 次の条件を付加した.

形式動詞, 形式名詞文節

予備実験では, 形式動詞「する, なる」, 形式名詞「こと...」機能動詞「与える, 受ける, 始める, かかる...」などからなる文節 (形式文節) の一致だけでは不正な編集が多数発生した. 形式文節はそれだけでは意味を持たないため, 一つのニュース記事の中であっても異なる事実を示すことがあるためである. 形式文節の意味は, 係っている格要素が決める. そこでトリガーの資格を検査するには, 係っている格要素の一致を評価するルールを作成した.

活用の同一視

活用語の場合, 完全一致せずともトリガーになれるものがある. しかし時制, 態, 法が違う場合には, 異なる事実を指すことからトリガーにはならない. この条件は分類しきれておらず, 今回は, 活用語の語尾「～し」と「～しました」のみを同一視することにした. 例えば, 表 1 の「会談し」と「会談しました」を同一視する.

4.2 最大句の対応付け処理

次に示す句対応率を使い, どん欲法でこの値の高い順に対応付ける. 対応率が閾値以下になれば対応付けをやめる.

句対応率 P : 句 \times 句 $\rightarrow [0, 1]$

X を最大句 X, Y のうちの文節数の少ない方とする. また x, y を最大句 X, Y の最終文節とする. 句対応率は次のように計算する. 本稿では経験的に $\alpha = 0.375$ とした.

$$P(X, Y) := \alpha D(X, Y) + (1 - \alpha) s(x, y), \quad (0 \leq \alpha \leq 1)$$

この式は句全体としての含有率 (第 1 項) と, 最終文節の付属語を含めた類似度 (第 2 項) の加重平均である. 最終文節を付属語まで含めて評価しているのは, これが句の文法的な性質を決める上で重要だからである.

4.3 挿入処理

挿入候補に対して, 「結束性」と「重複率」の検査を行い, これを満たせば, 挿入位置を決定して候補として提示する.

4.3.1 結束性検査

例えば指示語で始まる句をリードに挿入すると, 多くの場合, 不自然となる. なぜなら指示語は先行詞を必要とするが, 記事の冒頭のリードに挿入すると, 先行詞が得られないからである. そこで, 指示語のように前の文脈を必要とする語を禁止語リストにまとめ, このリストにある語で始まる句の挿入を禁止した. このリストは, 「こそあど」や「接続詞」のほか, 「その後」「今回」「他の」「同じ」などを含んでいる.

4.3.2 重複検査

挿入しようとする最大句はリードのトリガー支配範囲の句とは類似していない. しかし, その外側に類似した句がリードにあると, 情報が重複する. そこで, 挿入する最大句とリードの重複度合いを次式で測り, 閾値を超えた場合は挿入を禁止した.

重複率 O : 句 \times リード文 $\rightarrow [0, 1]$

挿入しようとする最大句を X , リード文を L として句の含有率を計算し, これを重複率とする. 句の含有率は文節連続を見ているだけなので文でも計算できる.

$$O(X, L) := D(X, L)$$

4.3.3 挿入位置決定処理

挿入句の右方向にある置換句 (リードに対応相手のある句) を探し, その対応先の左を挿入位置とした. 置換句がなければ, トリガーの左とする. これはソース文での句の順序を反映させるためである. 図 1 の「4 日,」は, 右隣の置換句の「IAEA の査察官...」の置換先, 「IAEA のチーム...」の左, すなわち先頭となる.

4.4 置換処理

置換候補の 2 つの対応する句中に, 「完全一致文節」があれば分割し, 次に具体性の検査を行う. その後, 挿入処理に準じて結束性検査と重複検査を行い, これらを満たせば置換候補として提示する. ここでは置換処理に特有な「句の分割」と「具体性検査」について説明する.

4.4.1 句の分割処理

2 つの対応句の中に完全一致する文節があれば, トリガーとこの間を置換候補とする. なければ対応句全体を置換候補とする. 図 1 の置換では「IAEA の」が完全一致文節のため「チームが」と「査察官 5 人が」が置換候補となる. 「IAEA の」の前に異なった句が係っていれば, 「IAEA の」をトリガーとした処理が別に行われる. 本処理は包含関係を持つ編集候補を出すのを避け, これを分割するのが目的である.

4.4.2 具体性検査

ソースの挿入候補句とリードの対応句の文節数を比較して, ソースの挿入候補の方が

多ければ挿入可能とする。文節数が同じ時は文字数で判定する。多くの場合、長い表現の方がより具体的であるという経験則による。

5. 主観評価実験

5.1 実験要領

提案手法は、経験的な知識に基づいて作成した規則を利用している。これらの有効性を調べるための実験を行った。

対象: 2004年1月20日, 4月20日, 7月20日のニュース記事257記事。各文はmecabとcabochaで解析してあり, 人手でリード, 本記, 追記のタグが付与されている。

評価者: 日本語母語話者1名。要約の専門家ではない。

評価インターフェース: 専用の評価インターフェースを作成した。これはリードとそれ以外の文(ソース)の間で発生するすべての置換と挿入を個々に提示する。置換では削除と挿入が同時に発生するため, 削除される表現を括弧に囲んで表示する。このため評価者は両方を見比べて評価することができる。

評価手順と評価項目: 評価者はそれぞれの編集結果が正しい係り受け解析から得られたかどうかをまず評価した。このとき, 文全体の係り受けでなく, トリガーに係る最大句など, 編集に関わる部分の係り受けのみを評価した。次に, 以下の項目を5段階で主観評価した。5段階評価は[9]に習い次のように設定した。評価項目の数値は, 結果をまとめるために与えた間隔尺度である。

評価項目は, 編集されたリードそのものを個別に評価する絶対評価と, 編集されたリード文を元のリードと比較する相対評価に大別される。個々の評価の狙いは以下の通りである。

絶対評価

意味的整合性: トリガー選択を誤ると, 同じ事実を表さない文節ペアに係る最大句を挿入・置換することになり, 意味の通らないリードとなる。すなわちこの項目ではトリガー選択の規則(4.1節)を主に評価する。なお, ここでは文法的適切性は評価していない。さらにこの項目が1点であった場合, 以下の項目はすべて1点とした。これは意味の通らない文に対する詳細な評価が難しかったからである。

重複度合い: 形態素の重複に基づいた重複検査規則(4.3.2節)の有効性を評価する

文法的適切性: 編集結果の係り受けの正しさを評価する。主に, 編集された句の助詞の適切性などを評価する

操作単位: 置換・挿入の最大句の長さの適切性を評価する。特に係り受け解析に失敗すると, 最大句の認定を誤ることがあり, このときには最大句が不正

表4 評価基準

大変よい	5
よい	4
かろうじて許容範囲	3
悪い	2
大変悪い	1

表5 各項目の平均点と統計的な有意差

状態	1) 正, 本記	2) 誤, 本記	3) 正, 追記	4) 誤, 追記
意味的整合性	4.66 (A)	4.48 (B)	2.44 (C)	2.33 (C)
重複度合い	4.31 (A)	3.72 (B)	2.42 (C)	2.17 (C)
文法的適切性	4.33 (A)	3.83 (B)	2.19 (C)	1.72 (C)
操作単位	4.52 (A)	3.43 (B)	2.49 (C)	2.06 (C)
指示明確性	4.55 (A)	4.57 (A)	2.56 (B)	2.33 (B)
挿入・置換位置	4.27 (A)	3.89 (B)	2.51 (C)	2.33 (C)
十分性	4.58 (A)	4.51 (A)	2.47 (B)	2.33 (B)
具体性改善	4.49 (A)	4.38 (A)	2.19 (B)	2.33 (B)
全体	3.77 (A)	3.01 (B)	1.83 (C)	1.39 (C)
頻度	378	169	59	18

となる。

指示明確性: 文脈を要求する語のリストを使った処理(4.3.1節)の効果を評価する

挿入・置換位置: 挿入位置決定規則(4.3.3節)を評価する。

十分性: 編集された結果は意味的に十分な情報を持っているかを評価する。中途半端に情報を追加したのでは意味的に不完全となる。操作単位が不適切なときには十分性が確保されないことがある。

相対評価

具体性改善: リードと比較して, 具体性がどの程度改善したかを評価する。表4の3点は, 「リードよりかろうじて具体性が増した」と考え, 他も同様とする。

全体: リードと比較した全体的な良さを評価する。表4の3点は「リードと比較してかろうじて良い」と考え, 他も同様とする。

5.2 結果

編集の例を示す。下線が挿入で () 内が削除された句である。1)は意味的にも, 文法的にも完全な例である。2)は文法的な問題はないものの, 具体性が向上しない例である。3)は, 具体性は向上したものの, 「から」が重なり文法的に問題な例である。

1) 民間団体の「コリア・ソサエティ」などが共催する「朝鮮半島平和フォーラム」

に(催しに)出席する

2) 部品に亀裂が入っているのが()見つかった

3) ヘリコプターから地上二十メートルの高さから () 落下し死亡しました。

評価結果を表5に示す。これは評価結果の平均点を係り受け状態(正, 誤), およびソース文の種類, (本記, 追記)の組の4状態でまとめたものである(第1行)。このとき, なるべく多くの評価項目の平均点が左から右に小さくなるように状態を並べた。このように並べた結果に対して, 隣のセルとの間に統計的な有意差 $[f]$ があれば異なるアルファベットを与えた。なお状態1(正, 本記)(係り受けが正しく, ソースが本記)には記号Aを与えている。すなわち, 評価項目を左から右に見たときに, アルファベットが変わらないところは統計的な有意差がないようになっている。

編集ソースの影響

状態2)と3)の間, すなわちソースが本記, 追記の境界をみると, すべての評価項目で大きな点差があり, これらに有意差が認められた。特に状態3)の全体評価が1.83とリード(3点を想定)より悪いことから, ソースを本記に限るのがよいことがわかる。

係り受けの影響

係り受け解析の誤り率は, 最下段の行より約30%(187/624)となった。係り受けの影響を, 本記(状態1と2)と追記(状態3と4)で個別に見ると, 追記では係り受けの正誤の影響はみとめられなかった。一方, 本記の中では9項目中6項目で有意差が認められた。これらの中で特に点差があったのが「重複度合い, 文法的適切性, 操作単位, 全体」であった。「全体」は相対評価なので除外すると, 文法的な側面の評価項目への影響が大きく, 意味的には影響が少なかったことがわかる。

全体の性能と課題

現状の性能については次のように考える。著者らは記事の各文に対して「リード, 本記, 追記」のタグを自動付与する研究を行っており, 92%程度の性能を達成している[8]。このため, 編集ソースを本記に限ることはできると仮定する。一方, 係り受け解析は現状程度と仮定する。この場合(表5の状態1と2を合わせた部分)の「全体」評価は3.54点である。オリジナルのリードの得点を3点付近と想定していることから, リードより改善されており, 手法の基本的な有効性は示せたと考える。特に, 文法性には問題が残るが, 具体性の増加や意味的整合性といった意味的側面が良かった。今後の課題として, まず係り受け解析の性能向上は必須と考える。ただし係り受け解析が正しくとも「全体」が3.77点であることから, 他の項目の改善も必要である。これには表5の第2列で4.5点以下である, 「重複度合い」「文法的適切性」「挿入・置換位置」があげられる。「重複度合い」「文法的適切性」については, 最大句を変更しな

f) ここでは順序尺度の検定手法である Mann-Whitney の U 検定を採用し, p-値を Fisher の正確確率を使って計算した。p-値 0.05 を有意差の判定基準とした。

いで挿入・置換する本手法の限界の可能性があり, 拡張が必要と思われる。例えば重複については, 重複率が閾値を超えた場合に編集を許可しないようにしているが, 閾値に近い重複は発生する。最大句中の重複箇所を削除できれば問題は解決できると思われる。同様に, 「文法的適切性」の典型的な問題である助詞の重なり, 不適切な利用についても, これらを修正することで解決可能である。「挿入・置換位置」については, 挿入・置換の後で係り受け状態が変わることが問題である。いずれの問題も文融合法で提案されているように text-to-text generation ととらえて最適に生成し直す方針が有効と思われる。

6. おわりに

リード文を他の文の表現で具体化する手法を報告した。本稿ではアルゴリズムを中心に報告したが, 著者らは, すでに, ニュース構造を認定するシステムと組み合わせた編集候補提示システムを作成している。これについては稿を改めて報告したい。

参考文献

- 1) H. P. Luhn. 1958. The Automatic Creation of Literature Abstracts. *Advances in Automatic Text Summarization. The MIT Press*: 15-21.
- 2) Inderjeet Mani, Barbara Gates, and Eric Bloedorn. 1999. Improving Summaries by Revising Them. *Proc. of the 37th Annual Meeting of the Association for Computational Linguistics.*: 558-565.
- 3) Jahna C. Otterbacher, Dragomir R. Radev, and Airong Luo 2002, Revisions that Improve Cohesion in Multi-document Summaries: A Preliminary Study. *Proc. of the ACL-02 Workshop on Automatic Summarization*: 27-36.
- 4) Ani Nenkova. 2008. Entity-driven Rewrite for Multi-document Summarization, *proc. of the 3rd International Joint Conference on Natural Language Generation*: 118-125.
- 5) Regina Barzilay and Kathleen R. McKeown. 2005. Sentence Fusion for Multidocument News Summarization. *Computational Linguistics*. 31(3): 298-327.
- 6) Katja Filippova and Michael Strube. 2008. Sentence Fusion via Dependency Graph Compression. *proc. of the EMNLP 2008*: 177-185
- 7) 田中他. 2005. ニュース要約の実態調査と要約モデルの検討, 自然言語処理研究会資料, 2005-NL-170: 115-120
- 8) 田中他. 2007 ニュース要約のための簡易文脈解析, 自然言語処理研究会資料, 2007-NL-182: 75-82
- 9) Hoa Trang Dang. 2006. Overview of DUC2006. <http://www-24.nist.gov/projects/duc/pubs/2006papers/duc2006.pdf>