

自動要約における誤り分析の枠組み

西川 仁^{1,a)}

概要: 本稿では自動要約の誤り分析の枠組みを提案する。本稿では要約器が作成する要約が満たすべき3つの要件と誤った要約が生じる5つの原因とを定義し、それに基づいて要約に含まれる何らかの誤りを15種類に分類する。本稿の提案する誤り分析の枠組みに基づいて要約器が作成した要約を分析し、その結果を報告する。さらに、分析の結果に基づいて要約器を改良し、改良の結果、要約の品質が大幅に改善することを示し、それを通じて誤り分析の枠組みの有効性を示す。

HITOSHI NISHIKAWA^{1,a)}

1. はじめに

本稿では自動要約の誤り分析を扱う。自動要約研究の題材として広く用いられるコーパスの多くは数十から数百の入力文書と参照要約の組からなる。多くの場合、自動要約の入出力はいずれも複数の文からなる文章であり、機械翻訳のように文ではなく、また多くの自然言語解析のように文と何らかの中間表現でもない。そのため、誤りの分析において考慮しなければならない要素が多く、数十といった入出力の規模でも、必ずしも ROUGE [3] などの定量的な評価指標による表層的な分析以上の分析が、十分になされているとはいえない。そのため、何らかの誤りを含むと思われる要約をどのように分析すればよいのかという体系的な方法論は存在せず、したがって自動要約分野の研究者が各々の方法論をもって分析を行っているのが現状と思われる。

本稿では、自動要約における誤り分析の枠組みを提案する。まず、要約器が作成する要約が満たすべき3つの要件を提案する。また、要約器がこれらの要件を満たせない原因を5つ提案する。3つの要件と5つの原因から、15種類の具体的な誤りが定義され、本稿では、自動要約における誤りはこれらのいずれかに分類される。

本稿の構成は以下の通りである。2節では本稿が置く基本的な前提について説明し、本稿での議論の範囲を明らかにする。3節では誤り分析の枠組みを提案し、自動要約の

誤りが提案する15種類の誤りのいずれかに分類できることを示す。4節では実際の要約例に含まれる誤りを提案した枠組みに基づいて分析した結果を示す。5節では4節で得られた分析の結果に基づいて要約器を改良し、要約の品質が改善することを示す。6節では関連研究について述べる。7節では本稿をまとめ、今後の展望について述べる。

2. 基本的な前提

一般に、誤りといえば、本来得られるべき何らかの正しい結果があるものの、それとは異なる、すなわち正しくない別の結果が得られた際にそれを指しているものと思われる。文書分類であれば与えられた文書を正しい分類先に分類できなかった際にそれを誤りということが出来る。そのため、何らかの正しい結果、すなわち正解が定まらなければ誤りも定めることができない。

自動要約においては、この正解（以下、参照要約と呼ぶ）をいささか一意に定めづらい。これは、自動要約に限らず、自然言語生成を目標とする課題に共通する問題であるが、自動要約課題において、複数の人間の作業者に参照要約の作成を依頼すると、作業者に与える指示にもよるものの、まったく同一の参照要約が作成されるということはまずない。そのため、ある参照要約を基準とした際には誤りとなる要約が、別の要約を基準とした際には誤りとならないことがある。

本稿では、この問題は脇に置く。すなわち、ある1つの参照要約が存在するとき、それと要約器が作成した要約（以下、便宜的にこれを機械要約と呼ぶ）とを比較し、その差分を誤りとする。すなわち、何か差分があれば誤りを含

¹ 日本電信電話株式会社 メディアインテリジェンス研究所
Media Intelligence Laboratories, Nippon Telegraph and Telephone Corporation

^{a)} nishikawa.hitoshi@lab.ntt.co.jp

むし、そうでなければ誤りを含まない。誤りについては次節にて述べる。この単純化は以下の理由に基づく：

- 単一の参照要約の誤り分析の枠組みが存在しない状況において、複数の参照要約の誤り分析の枠組みを設定するのは困難であると考えられること。
- 単一の参照要約の誤り分析の枠組みを設定できれば、それに基づいて複数の参照要約が存在する場合を検討することができると思われること。

これらの点から、本稿でのこの単純化は、問題の過度な単純化ではなく、合理的な問題の分割であると考えられる。

3. 誤り分析の枠組み

ここではまず、自動要約が最低限満たすべき原則を3つ述べ、それが満たされないときに誤りが生じることを説明する。次に、誤りの原因を5つ取り上げる。最後に、これから要約の誤りが15種類に分類されることをみる。

3.1 自動要約の誤りの種類

我々は、要約器によって生成される要約は以下の3つの原則を満たすべきと考える：

- (1) 出力から情報を読み取れること。情報を読み取れないような文章が出力されていないこと。情報を読み取れないような文が出力された場合には、以下の3つのケースが考えられる。
 - (a) 要約がユーザの要求とは異なる言語で出力されている場合や、要約器がその内部処理において利用している制御記号などが出力されており、要約から文意を読み取れない場合。何らかの理由により要約が出力されない場合も含む。
 - (b) 文法的でない文（非文）が要約を構成しており、要約の文意が取れない場合。
 - (c) 個別の文は文法的であるが、要約を構成する文同士の論理関係などが明らかでなく、全体として文意が取れない文章が要約となっている場合。

本稿ではこれら3点をまとめて、内容を適切に読み取ることのできない要約を便宜的に「非文章」と呼ぶ。

- (2) 読み取れる情報が、入力と矛盾せず、入力が入力を含意すること。読み手が入力を読んだ際と出力を読んだ際に異なる結論に至らないこと。
- (3) 出力から読み取れる情報が、入力および読み手の希望を鑑みて、重要であると思われること。重要でない、枝葉末節の情報が出力に含まれないこと。この点は自動要約の内容性評価と概ね対応する。

これらの原則から、したがって、これらの原則が満たされない場合を誤りとして、自動要約の誤りの分析における3つの観点が導出できる：

- (1) 非文章の出力：要約器が出力した文章から文意が読み取れない場合、それは誤りとなる。この観点は自動要

約の可読性評価 [4], [5] と概ね対応する。

- (2) 文意の歪曲：要約から読み取れる情報が、入力文書に記載されている情報と矛盾する場合、それは誤りとなる。この観点はこれまで自動要約において大きく取り上げられてこなかった。これには2つの理由が考えられる。第1に、現時点では、この観点に関して機械要約を評価するためには人手での丁寧な読解が不可欠であり、そのため非常に費用がかかり実施しづらいということが挙げられる。上で述べた (1) については出力された機械要約のみを人手で確認すればよく、また次に述べる (3) については参照要約と機械要約の機械的な比較によって人手をかけずに評価が可能である。これらに対して、(2) を評価するためには入力文書と機械要約をいずれとも評価者が読解した上で、内容の無矛盾を確認しなければならず、その費用は多大なものとなる。第2に、文の書き換えなどを行わずに単に重要文を選択するだけの手法などで要約を作成した場合、文意の歪曲はさほど頻繁には生じず、そのため誤りとしてこれまで重要視されてこなかったということが考えられる。

- (3) 重要部同定の失敗：要約から読み取れる情報の中に入力文書および読み手の希望を鑑みて重要でないものが混ざっているとき、それは誤りとなる。同様に、入力文書および読み手の希望を鑑みて重要であると思われる情報が要約に含まれていない場合もそれは誤りとなる。この観点は内容性の評価に概ね対応する [5] 。

この3つの観点が、要約器の誤りを考える際に、最初の分類としてあらわれるものと思われる。

3.2 要約器の誤りの原因

一方、要約器が前節のように誤る原因には以下の観点が考えられる：

- (1) 操作の不足：要約器が、人間の作業者がテキストに対して施す操作と同等の機構を保持してないことに伴って生じる誤り。言い換えなどの操作ができないために入力された文を短縮することができず、人間と同等の情報量を要約に含めることができない場合や、要約器が入力された文において省略されているゼロ代名詞を復元できず、要約の文意を損なう場合が含まれる。
- (2) 特徴量の不足：特徴量が不足している場合。この場合は2つに分けることができる。
 - (a) 特徴量の設定不足：要約器において設定されていない特徴量が要約の作成において重要な役割を果たすと思われる場合。段落に関する情報を入力文書から得ることができ、かつその情報が要約の作成において重要な役割を果たすと目されるのにもかかわらず、要約器はそれを特徴量として認識できない場合など。

(b) 言語解析の失敗：解析器が誤り、特徴量として設定されている情報が正しく取得できなかった場合。固有表現認識器が固有表現を認識し損ね、要約器がそれを特徴量として利用できなかった場合など。

(3) パラメタの誤り：訓練事例の不足、不適切な学習手法の利用などによって、推定されたパラメタが十分でない場合。

(4) 探索の誤り：探索誤りのために誤った要約を生成した場合。重要文集合の選択において、本来はより良好な文の組み合わせがあるにもかかわらず、探索誤りによって不適切な文の集合を出力として選択した場合など。

(5) 情報の不足：そもそも要約器に対して入力された情報だけでは参照要約まで到達できない場合。人間の要約作成者が入力以外の情報源を利用して要約を作成した場合など。

6 節で述べるが、これらの誤りの原因はより詳細化することが可能である。一方、自動要約には単一文書要約と複数文書要約といういささか風合いの異なる2つの小課題が存在し、また文短縮なども独立した課題として扱っている。そのため、要約器の設計は様々であり、誤りの原因の詳細は分析の対象とする要約器の設計に依存する。このことを鑑み、本稿ではより詳細な誤りの原因には踏み込まず、多くの要約器において共通する誤りの原因として上の5種類の原因を定義する。

3.3 自動要約の誤り分析の枠組み

3.1 節で述べた3種類の誤りの種類と、3.2 節で述べた5種類の誤りの原因から、自動要約における誤りは15種類のいずれかに分類できると期待できる。これをまとめたものを表1に示す。

なお、これらとは別に、参照要約作成者の読みが誤っていると思われる場合など、そもそも参照要約が信頼できなと思われる場合がありうるが、ここではそれは除外し、あくまで参照要約が正しく、機械はそれを模倣することのみを考えればよいという場合を想定した。

次に、分析の枠組みを自動要約の結果に適用する際の具体的な方法について表2に示す。表2は、ある誤りの種類がある誤りの原因によって生じる際に、どのようにそれを同定できるかをまとめたものである。

4. 分析の実践

本節では前節で提示した分析の枠組みを、本稿で分析の対象とした文書に対して適用する。

まず、分析の枠組みの適用の対象とする機械要約を作成する。次に、それらに対して人手による分析を行い、その後分析の結果を提案した分析の枠組みに基づいて整理する。

表3 入力文書および参照要約、機械要約の統計量。

	文数	文字数
入力文書	32	1215
参照要約	13	495
機械要約	11	493

4.1 実験設定

4.1.1 データ

TSC-2^{*1} のフォーマル・ランのデータを用いた。その中でも作成者1による自由記述の要約を参照要約として取り上げ、特に、文書番号990305053を用いた。

4.1.2 要約器

要約器については、西川らによる単一文書要約器 [6] を利用した。文短縮は用いずに利用した。

4.2 結果

表4に入力文書(文書番号990305053)を示す。太字は入力文書と参照要約とで文アライメントを取り、対応づけが取れた文同士において共通の単語である。下線は要約器によって重要文と認定された文である。表5に参照要約を示す。分析の対象となると思われる点については下線を加え、どのような現象が生じているか下線の後に上付き文字で示した。表6に機械要約を示す。太字は参照要約と機械要約とで文アライメントを取り、対応づけが取れた文同士において共通の単語である。表5と同様に分析の対象となると思われる点について下線を加え、どのような現象が生じているか下線で示された部分の後に加筆した。表3に入力文書および参照要約、機械要約の統計量を示しておく。

4.3 誤り分析

4.3.1 重要部の同定の失敗

まず、ROUGE-1 [3] の値は0.385であった。文単位でみると、機械要約に含まれる文のうち、完全に参照要約に含まれない文は2文めと11文のみであり、11文中2文にとどまっている。このことから、要約器の精度(適合率)は $\frac{9}{11}$ に達しており、要約器は高精度に重要文を同定していることがわかる。一方、再現率の観点から見ると、参照要約は入力文書33文のうち15文を要約として採用しており^{*2}、再現率は $\frac{9}{15}$ に留まっている。このことから、単語単位での再現率の指標であるROUGE-1の値は、まだまだ改善の余地があることがわかる。

次に、重要部同定の失敗の原因を探る。表4を見ると、要約器は特に後半の文を選択できていない。これは、要約器が入力文書における話題の遷移を捕捉できていないためであると思われる。入力文書において、どのような話題が遷

^{*1} <http://lr-www.pi.titech.ac.jp/tsc/tsc2.html>

^{*2} 2つの文を1つの文としてまとめているケースがあり、そのため参照要約は13文から構成されている。詳しくは文融合の節にて詳述。

表 7 入力文書に含まれる話題の遷移。

話題番号	文	話題
1	1-2	全人代の開催
2	3-4	朱首相の中国の改革に対する決意
3	5-10	中国の改革に対する熱気の薄れ
4	11-16	中国に対する信頼を揺らぎ
5	17-21	金融改革における外資の取り扱い
6	22-31	香港に対する官僚的な対応
7	32	記事のまとめ

移しているかを表 7 に示す。全人代が開催されるということ（話題 1）と中国の改革とその行く末が危ぶまれるということ（話題 2-4）と、その具体的な例（話題 5-6）が並び、最後の文は入力文書のまとめとなっている。参照要約を見ると、参照要約の作成者はできる限りこれらの情報を網羅的に要約に含めることを狙っていることが読み取れる。要約器が後半の文を選択できなかったのはこのような話題の構造を理解することができなかったため、この構造を要約器に理解させることは重要部の同定に決定的に重要である*3。

4.3.2 括弧の除去

表 4 の例において頻繁に行われている操作の 1 つは括弧の除去である。括弧を通じて提供されている補足的な情報は全て要約から除去されていることがわかる。これによって文を短くし文字数を稼ぐことができるため、要約器もこの操作を実行できるようにする必要がある。

4.3.3 文短縮・言い換え

表 4 を見ると、文書全体にわたって文の書き換えが行われていることがわかる。不要な修飾などを除去する操作は文短縮あるいは文圧縮と呼ばれており [5]、この表 4 の例でも文 1、文 10 などと典型的に行われている。一方、文短縮は典型的には係り受け木の枝刈りを通じて行われるが、参照要約に含まれる文のうち係り受け木の枝刈りによって実現できるものは少数であり、参照要約作成者はより洗練された言い換え操作を通じて参照要約を作成していることがわかる。

4.3.4 文の融合

異なる複数の文から 1 つの文を作成することは文融合と呼ばれている [1]。参照要約を見ると、この操作が行われていることがわかる。表 8 から 11 にその例を示す。参照要約の中では 4 回この操作が行われており、入力文書における表現と比べ情報量を維持したまま文字数の削減が行われている。これらの操作によって削減された文字数を利用して参照要約作成者はさらに情報を要約に詰め込んでお

*3 なお、西川らの要約器ではこのような話題の遷移を段落を通じて獲得しているが、今回利用したデータでは段落に関する情報が失われているため、これを利用できなかった。

表 8 文融合の例 1。文頭の数字はそれぞれ入力文書および参照要約中の文番号である。

入力文書	(5) 本来なら改革 2 年目の今年が正念場となるはずである。(6) ところが現実には、改革の熱気は薄い。
参照要約	(4) 本来なら改革 2 年目の今年が正念場となるはずだが、現実には改革の熱意は薄い。

表 9 文融合の例 2。文頭の数字はそれぞれ入力文書および参照要約中の文番号である。

入力文書	(12) 例えば、朱首相が昨年公約した「8%成長の確保」は、7・8%に終わった。(14) だが西側の経済専門家からは「本当は7・8%より低いのではないか」という疑問が出されている。
参照要約	(8) 公約の8%成長は7・8%だったが、本当はこの数字より低いのではないかという疑問が専門家からも出されている。

表 10 文融合の例 3。文頭の数字はそれぞれ入力文書および参照要約中の文番号である。

入力文書	(17) 金融改革については、外資の取り扱いで大きく揺れている。(22) 香港に対しても、最近の中国の姿勢は、硬直した感じが否めない。
参照要約	(11) 金融改革も、外貨の扱いで大きく揺れており、香港に対する姿勢も、硬直した感じが否めない。

表 11 文融合の例 4。文頭の数字はそれぞれ入力文書および参照要約中の文番号である。

入力文書	(30) 香港の繁栄回復が、中国の改革と切り離せないことを肝に銘じているのは中国のほうだ。(31) にもかかわらず、中国の対応はあまりにも官僚主義的だった。
参照要約	(12) 香港の繁栄が中国改革と切り離せないことが、分かっているはずなのに、中国の対応は官僚主義的である。

り、この操作を行う機構を持たない要約器は再現率において劣後せざるを得ない。

4.3.5 省略

便宜的に「省略」としたが、「この」や「など」の表現を用いて、入力文書における情報を除去している箇所がある。表 12 に示す参照要約の文 3 では、朱首相の「三つの実行」のうち金融機構改革が失われており、これが「など」として表現されている。また表 13 に示す参照要約の文 6 では、改革と安定追求のジレンマを「この」で表現しており、同様に文字数を稼いでいる。

4.3.6 参照要約の信頼性

一方、参照要約の品質が疑われる部分もある。入力文書

表 12 省略の例 1. 文頭の数字はそれぞれ入力文書および参照要約中の文番号である。

入力文書	(4) 国有企業改革、行政機構改革は計二千数百万人規模の大リストラ計画であり、「命をかけてやる」と言い切った首相の強い決意に称賛の声があがった。
参照要約	(3) 国有企業改革や行政機構改革などを「命をかけてやる」と言い切った首相の決意に称賛の声があがった。

表 13 省略の例 2. 文頭の数字はそれぞれ入力文書および参照要約中の文番号である。

入力文書	(10) 安定追求とのジレンマがあっても意志の強いことで知られる朱首相は改革路線を貫くと期待したい。
参照要約	(6) このジレンマがあっても、朱首相は改革路線を貫くと期待したい。

の文 14 と文 15 とは並列の関係にはないと思われるため、参照要約の文 9 先頭の接続詞「また」は要約作成者の読みの誤りを示唆している。

4.4 誤り分析の枠組みの適用

ここまでの分析を、本稿で提案した誤り分析の枠組みに適用した結果を表 14 に示す。表 14 に示されているように、今回は文短縮などの書き換え機構を利用していないため、非文が出力されることはなかった。一方で、文を短く書き換える操作を行えないため、情報の被覆において参照要約に大きく劣後しており、これが低い再現率の直接の原因となっている。

5. 分析に基づく要約器の改良

本節では、4 節で述べた分析に基づいて実際に要約器を改良した結果について述べる。

5.1 文の書き換え操作の追加

表 14 に示したように、今回の事例において操作の不足は深刻な問題である。そのため、参照要約において行われている書き換え操作の一部を要約器も行えるようにした。

5.1.1 括弧の除去

西川らの要約器は括弧を除去する機能を持つ^{*4}ため、この機能を動作させるようにした。

5.1.2 文短縮

同様に、文短縮機能も動作させるようにした。

5.1.3 文融合

文融合を行う機能を西川らの要約器は持たないため、表 8 から 11 に示した文融合が行われた文を手で作成し、要

^{*4} 正確には、文選択の際に、入力文書に含まれる元の文とは別に、括弧を除去した新しい文を生成し、それも選択の候補に含まれるようになっていく。

約器が選択可能な文集合に加えた。

5.1.4 省略

文融合と同様に、省略が行われている文についても人手で参照要約と同様の文を作成し、それを要約器が選択可能な文集合に加えた。具体的には、表 12 および 13 の参照要約の文を入力文書の文の書き換え後の文として要約器に追加した。

5.2 特徴量の追加

表 14 に示したように、一部の特徴量を要約器が認識できないことは要約の作成に悪影響を与えている。そのため、分析の結果として重要と思われた特徴量を追加した。

5.2.1 段落情報に関する特徴量

4.3.1 節で述べたように、重要文の同定に失敗した主因の 1 つは入力文書の話題の遷移を捉えることができないためであった。西川らの要約器は段落に関する情報を特徴量として利用することができるため、入力文書に表 7 に基づいて段落情報を付与した。具体的には、同一の話題番号に属する文は同一の段落に属するものとした。西川らの要約器は段落の先頭の文を重要文として選択する傾向があるため、これによって各話題の先頭の文を重要文として選択できると期待できる。

5.2.2 最後の文に関する特徴量

表 5 の参照要約を見ると、入力文書の最後の文を入力文書におけるある種のまとめとして重要文とみなしていることがわかる。この点を鑑み、最後の文にはその文が最後の文であるとわかる特徴量を追加した。

5.3 パラメタの調整

最後に、パラメタの調整を手で行った。パラメタの調整は、調整後に要約器が生成する要約が参照要約に近づくように人手で各特徴量の重みを調整することで行った。具体的に行ったのは以下の調整である：

- 括弧が含まれる文の重要度を下げたようにした。参照要約においては入力文書に含まれる括弧は全て除去されているため、これが除去されるようにした。
- 冒頭の段落に含まれる文の重要度を下げたようにした。通常、新聞記事は逆三角形と呼ばれる構造をなしており [8]、冒頭の段落がほぼ当該記事の要約をなしている。そのため、西川らの要約器は冒頭の段落に含まれる文に大きな重みを与えている。しかし、今回分析の対象とした入力文書はいささか散文的であり、その点を鑑みてか参照要約の作成者は記事の冒頭以外からも多く文を選択している。このことから、冒頭の段落に含まれる文の重みを小さくし、文書全体から文が選ばれるようにした。
- 長い文が選ばれづらくなるようにした。参照要約は長い文をあまり含んでおらず、文短縮や文融合、省略が

表 18 要約器の改良による ROUGE の変化. Rw は書き換え操作を追加したこと, Ft は特徴量を追加したこと, Pm1 はパラメータを手で調整したことをそれぞれ示す. Δ で示した数値はある改良によってどの程度 ROUGE-1 の値が改善したかを示す.

	ROUGE-1
改良前	0.385
Rw	0.436 (Δ 0.051)
Rw + Ft	0.521 (Δ 0.085)
Rw + Ft + Pm	0.667 (Δ 0.146)

施された短い文を含んでいる. そのため, それらの文が選ばれやすくなるように文の長さに対して負の重みを与えた.

- 百分率の固有表現を含む文が選ばれやすくなった. 参照要約には中国の経済成長に関する具体的な百分率が含まれており, これらの情報が要約に含まれるように百分率の固有表現の重みを大きくした.
- 類似する文が選ばれづらくした. 西川らの要約器は文同士の類似度を特徴量として設定しており, 類似した文が要約に選択されやすくなっている. しかし, 今回分析の対象とした入力文書の参照要約を見る限り, 参照要約の作成者はできるだけ幅広い話題を入力文書において網羅しようとしているように観察される. そのため, むしろ類似する文は要約に含まれないようにした方がよいと思われたため, 類似する文が選ばれづらくするようにした.
- 段落の先頭の文の重みを大きくした. 5.2.1 節で述べたように, 参照要約の作成者は入力文書に含まれる様々な話題を網羅するように要約を作成したように思われる. 特に, 各話題に関する段落の先頭の文を参照要約の作成者は参照要約に含ませているように観察されるため, これらが要約に含まれやすくなるようにした.
- 最後の文の重みに大きな値を与えた. 5.2.2 節で述べた特徴量は新しく追加したものであるため, 当該特徴量に対する重みがパラメータ集合内には存在しない. そのため, 最後の文が選ばれるように最後の文であることを示す特徴量に大きな重みを与えた.

5.4 結果と考察

書き換え操作を追加したのちの機械要約を表 15 に示す. 書き換え操作および特徴量を追加したのちの機械要約を表 16 に示す. 書き換え操作, 特徴量およびパラメータ調整を追加したのちの機械要約を表 17 に示す. これらの要約器の改良による ROUGE の変化を表 18 に示す. Rw は書き換え操作が追加された要約の評価, Rw + Ft は書き換え操作および特徴量が追加された要約の評価, Rw + Ft + Pm は書き換え操作, 特徴量, およびパラメータ調整が追加された要約の評価である.

書き換え操作の追加によっていくらか ROUGE が改善されたものの, 表 15 が示すように, 書き換え後の文の一部は要約器によって選択されておらず, その効果が十分に発揮されていない. そのため, ROUGE の改善も必ずしも大きなものではない. このことから, 単に書き換え操作を追加するだけではなく, 書き換え後の文が重要文として選択されるように特徴量およびパラメータを調整しないといけないことがわかる.

次に, 特徴量の追加による影響についてみる. 表 18 が示すように, 特徴量の追加により, 大きく ROUGE が改善していることがわかる. これは全て段落情報に関する特徴量の影響である. 最後の文に関する特徴量は新しく追加したものであるため, この時点では生成される要約に対して影響を与えない. 参照要約の作成者は入力文書に含まれる各話題からそれらに対応する文を選択しているため, 段落情報を通じてこの情報を要約器が利用できるようになった影響は大きい.

最後に, パラメータの調整による影響をみる. 表 18 が示すように, パラメータの調整により ROUGE が劇的に改善していることがわかる. 表 17 に示す要約には参照要約に含まれていない文が 1 つだけ含まれているものの (文 12), 参照要約にかなり類似した要約を生成することに成功している. このことから, 適切な書き換え操作と特徴量を追加した上で適切なパラメータを得ることができれば, 参照要約に近い要約を生成できることがわかる.

6. 関連研究

ここでは, 自動要約と同様に自然言語を生成する課題として機械翻訳を, また自動要約とは異なり自然言語を解析する課題として語義曖昧性解消を取り上げ, それぞれ本稿で取り扱った自動要約の誤り分析と比較する.

まず, 赤部らによる機械翻訳の誤り分析 [9] を取り上げる. 機械翻訳は自動要約と同様にテキストを入力としてテキストを出力する課題であり, 誤り分析の形態も似通ったものになると期待される. 赤部らは誤り分析を 2 種類に分類している. 1 つはブラックボックス分析であり, システムの出力にのみ着目して誤りを分析するものである. もう 1 つはグラスボックス分析であり, システムの動作過程も分析の対象としたものである. 本稿で扱った誤り分析はシステムの動作過程と出力の両方を扱っていることから, グラスボックス分析に相当する. 本稿の 3.1 節で提案した誤りの種類のみ注目して誤り分析を行うのであればこれはブラックボックス分析になる.

赤部らの提案しているブラックボックス分析の誤り体系は, 出力のみを分析するものであり, その点において本稿の 3.1 節と概ね対応している. 本稿の提案した要約の誤りの種類は赤部らのブラックボックス分析の誤り体系を抽象化したものになっている. 例えば, 自動要約の誤りの種類

の2つめ「入力が出力を含意しない」の原因の1つとして赤部らのブラックボックス分析の誤り体系の「モダリティ」を考慮することができる。自動要約の満たすべき要件を敷衍し機械翻訳の誤りを考えると、「出力から(目標言語で)情報が読み取れること」「(言語は異なるものの)入出力が意味的に等価であること」の2点を要件として考えることができ、その点において赤部らの提案したブラックボックス分析の誤り体系の一部は自動要約の誤り分析のより具体的な誤りの分類として考えることもできよう。

本稿の提案した要約の誤りの種類と赤部らのブラックボックス分析の誤りの体系を比較すると、自動要約と機械翻訳には2つの違いがあることがわかる。1つは非文章の存在である。自動要約の出力は多くの場合、文ではなくて文章であるため、文としては解釈できても文章としては適切に解釈できない場合が生じる。一方、現在の機械翻訳は基本的には文単位の処理を行っている^{*5}。もう1つは、自動要約の満たすべき要件の3つめ「入力および読み手の希望を鑑みて重要な情報のみが出力に含まれること」という点である。自動要約はその名の通り、重要な情報のみを読み手に提示することが目標であるが、機械翻訳は入力を目標の言語に入力と意味的に等価に変換することが目標であり、重要な情報を選別するという要件が存在しない。

赤部らが提案したもう1つの誤り体系である、グラスボックス分析の誤り体系は本稿の3.2節で提案した要約の誤りの原因にほぼ直接対応している。対応を表19に示す。表19に示すように、自動要約の誤りの原因と赤部らのグラスボックス分析の誤り体系はほぼ直接対応している。これは、現在の自動要約器も機械翻訳器も、自然言語の入力に形態素解析器などの基本的な解析器を用いて適切な解析を加える機構、入力を入力とは異なる表現に変換する機構、変換された表現の中で正しいと思われるものに高いスコアを与える機構、高いスコアが与えられる表現を探索する機構の4点をその基盤としているためである。

次に、自然言語の解析を目的とする課題として語義曖昧性解消課題の誤り解析を取り上げる。新納らは7名の分析者による誤り分析の結果を統合し、語義曖昧性解消課題において生じる誤りの原因を9種類に分類している[10]。語義曖昧性課題における誤りは正しい語義に単語を分類することができないがために生じるものであり、その点において本稿で提案した自動要約の誤りの種類や赤部らのブラックボックス分析の誤りの体系のように複数の誤りの種類は存在せず、単に誤分類のみが誤りとなっている。

新納らの提案した9種類の誤り原因は、本稿で提案した5種類の誤りの原因の一部を詳細化したものとみなせる。対応を表20に示す。語義曖昧性解消課題は自然言語の生成を行わないため、当然、書き換え操作の不足に対応する

誤りは存在しない。同様に、候補となる語義のいずれかに単語を分類する問題であるため、複雑な探索も行う必要がなく、そのため探索の誤りも存在しない。

7. おわりに

本稿では、自動要約の誤り分析を扱った。自動要約の誤りの分類を提案し、それを利用して一つの文書の分析結果を分類した。また、どのような誤りが生じているかを調査するための具体的な方法についても提案した。それらを用いて、ある文書がある要約器を用いて要約したとき、内部でどのような誤りが生じているか分析した。さらに、分析の結果を踏まえて要約器に改良を施し、その結果を報告した。

今後は、提案した分類をより精緻化し、個別の分析事例を蓄積していく予定である。特に、「操作の不足」の内実、すなわちどのような操作が不足しているために自動要約が失敗しているかについてより詳細な分析を行っていく予定である。

謝辞 本稿は自然言語処理における誤り分析プロジェクト Project Next^{*6}の一環として行われた研究に基づくものである。その過程において、国立国語研究所浅原正幸准教授、東京工業大学奥村学教授、東京工業大学菊池悠太氏、早稲田大学酒井哲也教授、九州工業大学嶋田和孝准教授、ニューヨーク大学関根聡研究准教授、東京工業大学高村大也准教授、日本電信電話株式会社平尾努研究主任、および京都大学森田一研究員よりご助言を頂戴した。記して感謝する。

参考文献

- [1] Barzilay, R. and McKeown, K. R.: Sentence Fusion for Multidocument News Summarization, *Computational Linguistics*, Vol. 31, No. 3, pp. 297-328 (2005).
- [2] Hardmeier, C., Nivre, J. and Tiedemann, J.: Document-Wide Decoding for Phrase-Based Statistical Machine Translation, *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL)*, pp. 1179-1190 (2012).
- [3] Lin, C.-Y.: ROUGE: A Package for Automatic Evaluation of Summaries, *Proceedings of ACL Workshop Text Summarization Branches Out*, pp. 74-81 (2004).
- [4] National Institute of Standards and Technology: The linguistic quality questions (2007). <http://www-nlpir.nist.gov/projects/duc/duc2007/quality-questions.txt>.
- [5] Nenkova, A. and McKeown, K.: *Automatic Summarization*, Now Publishers (2011).
- [6] Nishikawa, H., Arita, K., Tanaka, K., Hirao, T., Makino, T. and Matsuo, Y.: Learning to Generate Coherent Summary with Discriminative Hidden Semi-Markov Model, *Proceedings of the 25th International Confer-*

^{*5} もちろん、文を越えた単位での翻訳の試みも存在する [2], [7] .

^{*6} <https://sites.google.com/site/projectnextnlp/>

- ence on *Computational Linguistics (Coling)*, pp. 1648–1659 (2014).
- [7] Xiong, D., Ding, Y., Zhang, M. and Tan, L. C.: Lexical Chain Based Cohesion Models for Document-Level Statistical Machine Translation, *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 1563–1573 (2013).
- [8] 一般社団法人共同通信社：記者ハンドブック新聞用字用語集，共同通信社，第12版 edition (2010).
- [9] 赤部晃一, Graham, N., 工藤 拓, John, R., 中澤敏明, 星野 翔：Project Next における機械翻訳の誤り分析，言語処理学会第19回年次大会ワークショップ「自然言語処理におけるエラー分析」発表論文集，pp. 1–5 (2015).
- [10] 新納浩幸, 白井清昭, 村田真樹, 福本文代, 藤田早苗, 佐々木稔, 古宮嘉那子, 乾 孝司：語義曖昧性解消の誤り分析，言語処理学会第19回年次大会ワークショップ「自然言語処理におけるエラー分析」発表論文集，pp. 1–25 (2015).

表 1 自動要約の誤り分析

		非文章の出力	文意の歪曲	重要部同定の失敗
操作の不足		<ul style="list-style-type: none"> ● 非文については、文を生成する、あるいは書き換える機構が不十分であるため非文が生成される場合など。 ● 非文章については、機械要約に適切な談話構造を与える機構が不足している場合など。 	<ul style="list-style-type: none"> ● ゼロ代名詞の復元を行う機構を要約器が備えていない場合。その場合、機械が作成した要約に対して、入力文書とは異なる理解がなされる可能性が生じる。 ● 文章を構成する論理関係が入力文書と異なる読みを許すものになっており、読者が誤った結論に到達する場合。 	<ul style="list-style-type: none"> ● 参照要約の作成者が行った操作を機械が再現することができず、そのため要約長の制約などから重要な情報を要約に含めることができなかった場合。例えば参照要約の作成者が略語化によって文字数を節約した場合、機械も同様の操作を行わない限り参照要約に到達できない。
特微量の不足	特微量の設定不足	<ul style="list-style-type: none"> ● 文の書き換え規則の適用に用いる特微量が得られず、誤って必須格の格要素を除去した場合など。 ● 入力文書の談話構造に関する情報がなく、出力する要約に適切な論理構造を与えることができない場合など。 	<ul style="list-style-type: none"> ● ゼロ代名詞を伴う文にそのことを表す特微量が付与されず、これを考慮せずに出力した要約が結果として文意を歪曲する場合など。 	<ul style="list-style-type: none"> ● 入力文書の特定期間が要約に含まれるべき重要な情報を含んでいることを、特微量の設定の不足によって機械が理解できない場合。固有表現や評価表現などの情報が付与されていない場合など。
	言語解析の失敗	<ul style="list-style-type: none"> ● 係り受け解析器が係り受け解析を誤った場合や、述語項構造解析器が述語項構造の解析に失敗した場合、談話構造解析器が談話構造の解析に失敗した場合など。これらが生じた場合、非文章が出力される恐れが高まる。 	<ul style="list-style-type: none"> ● ゼロ代名詞を述語項構造解析器が誤って復元した場合などが該当。入力文書の内容を要約が含意せず、致命的な誤りとなる。 	<ul style="list-style-type: none"> ● 自然言語解析の失敗によって適切な特微量を機械が取得できなかった場合。固有表現認識に失敗した場合など。
パラメタの誤り		<ul style="list-style-type: none"> ● 文の書き換え規則の適用順序が正しくなく、誤って必須格の格要素を削除してしまった場合などが該当。 	<ul style="list-style-type: none"> ● ゼロ代名詞を含む文とその先行詞を含むと思われる文があり、後者の文の重みが低く見積もられている場合など。 	<ul style="list-style-type: none"> ● ある特微量が適切な重みを得ておらず、重要文として認定されるべき文がそう認定されなかった場合。訓練事例の不足や、不適切な学習方法が用いられた場合などが含まれる。
探索の誤り		<ul style="list-style-type: none"> ● パラメタは問題がないが、最適解が得られなかったために文の書き換えに失敗した場合。貪欲法などの近似解法を用い、最適解に到達できなかった場合が含まれる。 	<ul style="list-style-type: none"> ● 左に同じ。 	<ul style="list-style-type: none"> ● 左に同じ。
情報の不足		<ul style="list-style-type: none"> ● 入力文書において言及される何らかの実体に関する情報が入力文書において十分に説明されておらず、出力となる要約を読んで読み手が要約を適切に解釈できない場合。 	<ul style="list-style-type: none"> ● 入力文書が本質的に曖昧性を含んでおり、外部の情報なしには入力を正しく解釈できない場合など。そのような場合、機械による解釈が誤り、結果として文意を歪曲した要約が作成される場合がありうる。 	<ul style="list-style-type: none"> ● 要約のために必要な情報がそもそも要約器に与えられていない場合。例えば、新聞記事のタイトルに含まれる情報を必ず要約に含めるように要約作成者が要約を作成しているにもかかわらず、要約器に対してはタイトルの情報が与えられない場合。

表 2 自動要約の誤り分析の枠組みの適用方法.

		非文章の出力	文意の歪曲	重要部同定の失敗
操作の不足		非文が含まれていないかを調査. 加えて, 入力文書と機械要約を比較し, 機械要約が読解できない原因を探る.	人手による入力文書および自動要約の読解. 含意認識器の利用も考えることができるが, 精度的に困難があるろう.	一部の入力文を人手によって参照要約に含まれるものと同じもの書き換えることによって, 当該操作を要約器が実行可能であった場合の重要部同定の性能を見積もることができる.
特徴量の不足	特徴量の設定不足	文の書き換え規則などを追加し要約結果の変化を調査.	解析器の結果の確認.	参照要約が含む重要な情報がどのような手がかりから得られるかを調査. 固有表現, 手がかり語など.
	言語解析の失敗	解析器の結果の確認.	左に同じ.	左に同じ.
パラメタの誤り		人手によるパラメタの調整と結果の目視. パラメタを正しく設定することによって要約器が正しく動作する場合は, どのようにすればそのようなパラメタを推定できるか逆算する.	左に同じ.	左に同じ.
探索の誤り		整数線形計画問題ソルバーを利用し, 最適解を得た上で, 要約器の出力と比較する.	左に同じ.	左に同じ.
情報の不足		要約に与えられた情報のみに基づいて入力文書を適切に解釈できるか人手で確認する. 外部の情報がない場合において, 人間でも適切な読みが不可能である場合, 機械でもそれは不可能である.	左に同じ.	入力文書のタイトルや, 挿入されている図表のキャプションなどが自動要約の際の重要な手掛かりとなっていないか確認する. それらをクエリとして与えクエリ依存要約とした場合の性能を調査する.

表 4 文書番号 990305053 のテキスト。毎日新聞'99 データ集より引用した。太字は入力文書と参照要約とで文アライメントを取り、対応づけが取れた文同士において共通の単語である。下線は要約器によって重要文と認定された文である。

- (1) 中国の国会、全国人民代表大会（全人代）が 5 日から始まる。
- (2) 朱鎔基首相の「政府活動報告」と予算案を審議し、私有制経済の存在を保障する憲法の一部改正などを行う予定だ。
- (3) 昨年 の全人代で、新首相に選ばれた朱首相は、「8%成長」と「三つの実行」（国有企業改革、金融体制改革、行政機構改革の3年以内解決）などを公約した。
- (4) 国有企業改革、行政機構改革 は計二千数百万人規模の大リストラ計画であり、「命をかけてやる」と言い切った首相の強い決意に称賛の声があがった。
- (5) 本来なら改革2年目の今年が正念場となるはず である。
- (6) ところが現実には、改革の熱気は薄い。
- (7) アジア金融危機の影響が中国 に及び、経済環境 が急速に悪化した。
- (8) 改革で生まれる失業者を他の産業に吸収できない。
- (9) 改革のテンポを緩めても、社会不安を抑え込むべき だという空気 が強まっている。
- (10) 安定追求とのジレンマがあっても 意志の強いことで知られる朱首相は改革路線を貫くと期待したい。
- (11) しかし、中国の経済 が悪化 するとともに、中国に対する信頼を揺るがせるような問題もいくつか発生している。
- (12) 例えば、朱首相が昨年公約した「8%成長の確保」は、7・8% に終わった。
- (13) ほぼ8%であり、公約は達成されたとされた。
- (14) だが 西側の経済 専門家からは「本当は 7・8% より低いのではないか」という疑問が出されている。
- (15) 電力消費量 や国内輸送量 が増え ていないのに、国内総生産（GDP）が増えるのはおかしいと統計の公正さに疑問が出された。
- (16) 広東省など 地方の成長率が10%を超えた のも水増しを疑われている。
- (17) 金融改革 については、外資の 取り扱いで大きく揺れて いる。
- (18) 昨年秋、突然倒産した広東国際信託投資公司（GITIC）の負債の処理について、「正規に登録された外資は返済を保証する」という中国政府の方針が、今年になって引っくり返った。
- (19) 外資は「貸手にも責任がある」と突き放された。
- (20) 各地方の国際信託投資公司（ITIC）にも同様の問題が飛び火している。
- (21) そこでも同じ方針が貫かれると、今後中国へ向かう勇気のある外資はなくなるかもしれない。
- (22) 香港に対 しても、最近の中国の姿勢 は、硬直した感じが否めない。
- (23) 香港人が中国国内でもうけた子供に香港居留権があると判断した香港の裁判所を、中国当局者が激しく批判した。
- (24) 「香港基本法」の解釈権は中国の全人代にある。
- (25) 一地方政府にすぎない香港の裁判所に解釈権はない、という趣旨だった。
- (26) 香港の最終審長官が、「全人代の解釈権を侵害する意図はない」と釈明して取捨された。
- (27) だが「1国2制度」に対する香港市民の自信はこの一件で急落した。
- (28) 中国の国有企業は、香港の株式市場で資金を調達する予定だった。
- (29) ところが香港の不況で、上場延期に追い込まれている。
- (30) 香港の繁栄 回復 が、中国 の改革 と切り離せないことを肝に銘じているのは中国のはずだ。
- (31) にもかかわらず、中国の対応は あまりにも 官僚主義的 だった。
- (32) 改革の直面する困難が大きければ大きいほど、柔軟な対応が必要になるだろう。

表 5 文書番号 990305053 の参照要約. 下線とその後の上付き文字はその部分においてどのような現象が生じているかを示したものである.

- (1) 中国の国会、全国人民代表大会が始まる。
- (2) 昨年、朱首相は「8%成長」と「三つの実行」などを公約した。
- (3) 国有企業改革や行政機構改革 など 省略 を「命をかけてやる」と言い切った首相の決意に称賛の声があがった。
- (4) 本来なら改革2年目の今年が正念場となるはずだが、現実には改革の熱意は薄い。文融合
- (5) アジア金融危機の影響で、中国でも経済環境が悪化し、社会不安を抑え込むべきという空気も強まっている。文融合
- (6) この 省略 ジレンマがあっても、朱首相は改革路線を貫くと期待したい。
- (7) しかし、経済悪化とともに、問題もいくつか発生している。
- (8) 公約の8%成長は7.8%だったが、本当はこの数字より低いのではないかという疑問が専門家からも出されている。
- (9) また 参照要約の信頼性、電力消費量が増えないのに、国内総生産が増えるのはおかしいと統計の公正さにも疑問が出された。
- (10) 地方の成長率が10%を超えたことも水増しを疑われている。
- (11) 金融改革も、外貨の扱いで大きく揺れており、香港に対する姿勢も、硬直した感じが否めない。
- (12) 香港の繁栄が中国改革と切り離せないことが、分かっているはずなのに、中国の対応は官僚主義的である。文融合
- (13) 改革の困難が大きければ大きいほど、柔軟な対応が必要だろう。

表 6 文書番号 990305053 の機械要約. 太字は太字は入力文書と参照要約とで文アライメントを取り、対応づけが取れた文同士において共通の単語である。下線とその後の上付き文字はその部分においてどのような現象が生じているかを示したものである。

- (1) 中国の国会、全国人民代表大会 (全人代)括弧の除去 が5日から始まる。
- (2) 朱鎔基首相の「政府活動報告」と予算案を審議し、私有制経済の存在を保障する憲法の一部改正などを行う予定だ。重要部同定の失敗
- (3) 昨年の全人代で、新首相に選ばれた朱首相は、「8%成長」と「三つの実行」(国有企業改革、金融体制改革、行政機構改革の3年以内解決)括弧の除去などを公約した。文短縮
- (4) 国有企業改革、行政機構改革は計二千数百万人規模の大リストラ計画であり、「命をかけてやる」と言い切った首相の強い決意に称賛の声があがった。
- (5) 本来なら改革2年目の今年が正念場となるはずである。
- (6) アジア金融危機の影響が中国に及び、経済環境が急速に悪化した。
- (7) 改革のテンポを緩めても、社会不安を抑え込むべきだという空気が強まっている。文短縮
- (8) 安定追求とのジレンマがあっても意志の強いことで知られる朱首相は改革路線を貫くと期待したい。文短縮
- (9) しかし、中国の経済が悪化するとともに、中国に対する信頼を揺るがせるような問題もいくつか発生している。文短縮
- (10) 例えば、朱首相が昨年公約した「8%成長の確保」は、7.8%に終わった。
- (11) 香港人が中国国内でもうけた子供に香港居留権があると判断した香港の裁判所を、中国当局者が激しく批判した。重要部同定の失敗

表 14 自動要約の誤り分析の一例.

		非文章の出力	重要部同定の失敗	文意の歪曲
操作の不足		文短縮を利用しておらず非文の出力は行っていない.	文短縮や文融合といった、要約作成者がテキストに対して施した操作を模倣する機構を要約器は持たず、結果として低い再現率に甘んじている.	文意の歪曲を招くような文の組み合わせは生じていない.
特徴量の不足	特徴量の設定不足	上に同じ.	外資の取り扱いに関する話題の認定および香港に関する話題の先頭となっている文の認定に必要な特徴量を保持していなかったことは直接的に重要部同定に悪影響を与えている. また、文書末の文の重要度を低く評価したのは、文書末の文をあらゆる特徴量を設定しなかったことによると考えられる.	上に同じ.
	言語解析の失敗	上に同じ.	解析結果を見る限りは解析結果の誤りによる特徴量の抽出の失敗は生じていない.	上に同じ.
パラメタの誤り		上に同じ.	機械要約では文書頭近辺の文が集中的に選択されているため、文書頭の文を選択しやすい傾向を改める必要がある. また、比較的短い文が参照要約には多く含まれているため、短い文が選ばれやすくなるようにする必要がある.	上に同じ.
探索の誤り		上に同じ.	利用している探索ルーチンは最適解を保証しており、探索誤りは生じない.	上に同じ.
情報の不足		上に同じ.	入力した情報だけで要約器は参照要約に近づけるように観察される.	上に同じ.

表 15 書き換え操作を追加したのちの文書番号 990305053 の機械要約.

<p>(1) 中国の国会、全国人民代表大会（全人代）が5日から始まる。</p> <p>(2) 昨年全人代で、新首相に選ばれた朱首相は、「8%成長」と「三つの実行」（国有企業改革、金融体制改革、行政機構改革の3年以内解決）などを公約した。</p> <p>(3) 国有企業改革、行政機構改革は計二千数百万人規模の大リストラ計画であり、「命をかけてやる」と言い切った首相の強い決意に称賛の声があがった。</p> <p>(4) 本来なら改革2年目の今年が正念場となるはずだが、現実には改革の熱意は薄い。</p> <p>(5) アジア金融危機の影響が中国に及び、経済環境が急速に悪化した。改革のテンポを緩めても、社会不安を抑え込むべきだという空気が強まっている。</p> <p>(6) 安定追求とのジレンマがあっても意志の強いことで知られる朱首相は改革路線を貫くと期待したい。</p> <p>(7) しかし、中国の経済が悪化するとともに、中国に対する信頼を揺るがせるような問題もいくつか発生している。</p> <p>(8) 例えば、朱首相が昨年公約した「8%成長の確保」は、7・8%に終わった。</p> <p>(9) 香港人が中国国内でもうけた子供に香港居留権があると判断した香港の裁判所を、中国当局者が激しく批判した。</p> <p>(10) 香港の最終審長官が、「全人代の解釈権を侵害する意図はない」と釈明して取捨された。</p>
--

表 16 書き換え操作および特徴量を追加したのちの文書番号 990305053 の機械要約.

- (1) 中国の国会、全国人民代表大会が5日から始まる。
- (2) 朱鎔基首相の「政府活動報告」と予算案を審議し、私有制経済の存在を保障する憲法の一部改正などを行う予定だ。
- (3) 昨年 of 全人代で、新首相に選ばれた朱首相は、「8%成長」と「三つの実行」(国有企業改革、金融体制改革、行政機構改革の3年以内解決)などを公約した。
- (4) 国有企業改革や行政機構改革などを「命をかけてやる」と言い切った首相の決意に称賛の声があがった。
- (5) 本来なら改革2年目の今年が正念場となるはずだが、現実には改革の熱意は薄い。
- (6) アジア金融危機の影響が中国に及び、経済環境が急速に悪化した。
- (7) 改革のテンポを緩めても、社会不安を抑え込むべきだという空気が強まっている。
- (8) 安定追求とのジレンマがあっても意志の強いことで知られる朱首相は改革路線を貫くと期待したい。
- (9) しかし、中国の経済が悪化するとともに、中国に対する信頼を揺るがせるような問題もいくつか発生している。
- (10) 香港人が中国国内でもうけた子供に香港居留権があると判断した香港の裁判所を、中国当局者が激しく批判した。
- (11) 香港の繁栄が中国改革と切り離せないことが、分かっているはずなのに、中国の対応は官僚主義的である。

表 17 書き換え操作, 特徴量およびパラメタ調整を追加したのちの文書番号 990305053 の機械要約.

- (1) 中国の国会、全国人民代表大会が5日から始まる。
- (2) 昨年 of 全人代で、新首相に選ばれた朱首相は、「8%成長」と「三つの実行」などを公約した。
- (3) 国有企業改革や行政機構改革などを「命をかけてやる」と言い切った首相の決意に称賛の声があがった。
- (4) 本来なら改革2年目の今年が正念場となるはずだが、現実には改革の熱意は薄い。
- (5) アジア金融危機の影響が中国に及び、経済環境が急速に悪化した。
- (6) 改革のテンポを緩めても、社会不安を抑え込むべきだという空気が強まっている。
- (7) このジレンマがあっても、朱首相は改革路線を貫くと期待したい。
- (8) しかし、中国の経済が悪化するとともに、中国に対する信頼を揺るがせるような問題もいくつか発生している。
- (9) 例えば、朱首相が昨年公約した「8%成長の確保」は、7・8%に終わった。
- (10) 広東省など地方の成長率が10%を超えたのも水増しを疑われている。
- (11) 金融改革も、外貨の扱いで大きく揺れており、香港に対する姿勢も、硬直した感じが否めない。
- (12) 香港人が中国国内でもうけた子供に香港居留権があると判断した香港の裁判所を、中国当局者が激しく批判した。
- (13) 改革の直面する困難が大きければ大きいほど、柔軟な対応が必要になるだろう。

表 19 自動要約の誤りの原因と機械翻訳のガラスボックス分析の誤り体系との対応.

機械翻訳	自動要約
前処理誤り	特徴量の不足, 特に言語解析の失敗に対応. 言語解析器が解析を誤り, 正しい特徴量を抽出できない場合に相当.
ルール抽出	操作の不足に相当. 適切な翻訳ルールを取得できないことによる誤りは適切な書き換え操作を取得できないことによる誤りと類似.
モデル化誤り	パラメタの誤り. 学習過程における誤りによって誤った要約候補に高いスコアを与えることと直接対応する.
探索誤り	探索の誤りと直接対応する.
原文の誤り	情報の不足に対応. 自動要約において入力文書の適切な解釈ができないことによる要約の失敗と対応する.

表 20 自動要約の誤りの原因と語義曖昧性解消の誤りの原因の対応.

語義曖昧性解消	自動要約
訓練データの不足	パラメタの誤りに対応. 訓練事例が不足しているために正しいパラメタが推定できなかった場合に相当する.
深い意味解析が必要	特徴量の不足, 特に特徴量の設定不足に対応. 深い意味解析を行う解析器が存在せず, そのため必要とされる特徴量が入手できない場合に相当する.
シソーラスの問題	特徴量の不足に対応. 解析に際して必要な情報がシソーラスから欠落しており, そのため必要とされる特徴量が入手できない場合に相当する.
学習アルゴリズムの問題	パラメタの誤りに対応. 不適切な学習アルゴリズムを用いたために正しいパラメタが推定できなかった場合に相当する.
構文・格・項構造の素性不足	特徴量の不足, 特に特徴量の設定不足に対応.
素性のコーディングが困難	特徴量の不足, 特に特徴量の設定不足に対応.
テスト文に問題あり	情報の不足に対応. 入力となるテスト文に手がかりがないため, 正しい語義を推定できない場合は, 自動要約において入力に本質的な不足があるために適切な要約を生成しようがなく不良設定問題となっている場合と対応する.
共起語の多義性	特徴量の不足, 特に特徴量の設定不足に対応する.
データの誤り	本稿では誤りに含めなかったが, 参照要約が誤っている場合に相当する.