

## アスペクト被覆を可能にした 最小値最大化問題に基づく文書要約モデル

牧野 拓哉<sup>†1</sup> 高村 大也<sup>†2</sup> 奥村 学<sup>†2</sup>

本稿では複数のアスペクトを網羅的に被覆した要約の生成を最小値最大化問題として実現する手法を提案する。アスペクトは、読者が元の文書の内容を知るために重要な項目であり、予め定義されている。例えば文書クラスタのトピックが事故や天災であれば、アスペクトは“原因”や“被害”などがある。この複数のアスペクトを被覆することで全体的な内容を要約という形で読者に提示することが本稿の目的である。各アスペクトを捉えるために、我々は最大エントロピー分類器を用いた。我々のモデルは、最大エントロピー分類器の予測結果をアスペクトの反映度と考えて最小値最大化問題を解く。アスペクトの反映度のうち、最も小さな反映度を最大化するように要約を生成するため、要約は網羅的にアスペクトを被覆する。アスペクトが定義されている要約タスクにおいて、アスペクトを明示的に考慮することで、要約で一般的な重要な要素のみを考えたモデルよりも ROUGE-2 が向上することを示す。

### Balanced Coverage of Aspects for Text Summarization

TAKUYA MAKINO,<sup>†1</sup> HIROYA TAKAMURA <sup>†2</sup>  
and MANABU OKUMURA <sup>†2</sup>

We present a new text summarization model based on max-min problem to cover aspects. Aspects are pre-defined for each category of source documents; for example, the category *Accidents and Natural disasters* has aspects such as *WHY* and *DAMAGES*. Our goal is to generate a summary that covers these aspects. In order to calculate the score indicating the aspect coverage, we use the maximum-entropy classifier that predicts whether each sentence reflects the aspect or not. In our model, the score indicating the coverage for each aspect is calculated and the minimum of the scores of the aspects is going to be maximized so that the summary contains all the aspects. Through the summarization experiments on the TAC dataset, we show that our model outperforms a state-of-the-art summarization system in terms of ROUGE-2.

### 1. はじめに

文書要約とは、与えられた単一あるいは複数の文書から、その内容を簡潔に表した短い文書（要約）を生成する研究課題である<sup>5)</sup>。読者がある事象についての内容を知りたい場合、例えばその事象に関する文書集合が大量に存在する状況であれば文書要約技術は重要な技術となりうる。文書要約タスクには、要約に含めるべき情報に指向性がある topic-biased な要約タスクや、要約に含めるべき情報に指向性がない generic な要約タスクがあるが、本稿では topic-biased な要約に焦点を当てる。topic-biased な要約タスクは Text Analysis Conference (TAC 2010, 2011) の評価型ワークショップで課題として採用されており、TAC では topic-biased な要約を guided summarization と呼んでいる。guided summarization では、予め複数のアスペクトが与えられており、すべてのアスペクトに関する情報を要約へ含める必要がある。本稿が提案する手法は TAC のタスクで配布されているデータを用いて実験および評価をおこなっているので、以降 TAC での呼称に従う。

本稿が取り組む guided summarization では、予め与えられた複数のアスペクトを網羅的に要約へ含める必要がある。アスペクトを要約へ含める手法として、貪欲にアスペクトを最も反映しているような文を選択して要約を生成する抽出型の要約はあるが<sup>2)</sup>、この手法では局所解に陥る可能性があり、必ずしも大域解を求められる保証がない。そこで、本稿では最小値最大化問題として要約を定式化することで最適な要約を生成する手法を提案する。最小値最大化問題は整数計画問題 (Integer Linear Programming Problem (ILP)) として表現することができ、大域的な視点で要約を生成できる。要約に含まれるアスペクトの反映度を予測し、それらの反映度の中で最小となるアスペクトの反映度を最大化することでアスペクトを網羅的に被覆した要約を生成する。

ILP に基づいた複数文書要約手法において、概念単位の重要度を決定するために、文書頻度や新聞記事の特性を活かした文書中の出現位置を利用することが有用であることは知られているが、<sup>1)</sup>guided summarization では、被覆すべき内容にアスペクトという指向性がある。そのため、文書頻度や文書中の出現位置を利用して概念単位の重要度を決定するのみでは必ずしもアスペクトを被覆できるとは限らない。例えば、文書クラスタのトピックが *Accidents*

<sup>†1</sup> 東京工業大学 総合理工学研究科

<sup>†2</sup> 東京工業大学 精密工学研究所

and Natural disasters であり、かつアスペクトのひとつが COUNTERMEASURES という救助内容などを表すアスペクトであるとき、救助内容に関する情報が多くの文書で述べられていない場合には必ずしも救助内容に関する情報を要約へ含められない。そこで明示的にアスペクトを考慮することで、文書頻度や文書中の出現位置のみに依らずにアスペクトを被覆した要約を生成できるという直感が本稿の動機である。

我々はアスペクトを捉えるために最大エントロピー分類器を用いた。最大エントロピー分類器は確率的に事例があるクラスに所属する条件付き確率を予測することができる。我々はその条件付き確率が文があるアスペクトを反映する度合いと考えた。最大エントロピー分類器を訓練するためには、Wikipedia を用いた。Wikipedia は構造化文書であり、すべての記事にはカテゴリがついており、また記事はセクション毎に段落が分かれている。我々はこれらの特徴を利用して訓練データを作成した。

## 2. 関連研究

### 2.1 整数計画問題 (ILP)

近年、文書要約の分野で ILP を利用した手法が盛んに研究されている。ILP は、変数の取りうる値が整数であるような制約がついた線形問題で表される。文書要約問題は長さの制約の範囲内で元の文書クラスタの情報を最大化するような問題として表現されることが多い。ILP は大域的な解が求まることが知られている。ここで要約技術の背景について簡単に述べておく。まず Goldstein<sup>3)</sup> は、文の逐次選択による文書要約を実現した。彼らは、すでに選択された文と似た文に対してペナルティを与えることで要約における冗長性を排除した。ここで使われるペナルティは maximal marginal relevance (MMR) と呼ばれる。McDonald<sup>4)</sup> は、文書要約を近似的な動的計画法として定式化し、ILP として文書要約を表現した。高村ら<sup>5)</sup> は、多様性に加えて与えられた文書クラスタの主題への関連性を考慮した文書要約モデルへの拡張をおこなっている。

### 2.2 Wikipedia を利用した要約生成方法

Wikipedia を利用した要約生成方法は複数ある。Vasudeva ら<sup>2)</sup> は、regression モデルを用いて文単位でアスペクトの反映度を予測し、最も大きい反映度を持つ文から順に選択して要約を生成した。藤井ら<sup>6)</sup> は、Wikipedia のセクションを利用して検索対象の用語について自動的に観点を決定し、すべての観点について Support Vector Machine (SVM) を用いて最もふさわしい文を選択している。

## 3. 提案手法

### 3.1 アスペクトの反映度の計算方法

アスペクトは、文書クラスタの内容を知る上で重要な項目であり、予め定義されている。また、アスペクトは複数存在するもので、文書クラスタのトピック<sup>\*1</sup>によって異なる。例えば、元の文書クラスタのトピックが、Accidents and Natural disasters であれば、アスペクトは WHY, WHO AFFECTED, HOW, COUNTERMEASURES などがある。このようなアスペクトは TAC で行われている guided summarization task に基づいている。以下に、TAC のガイドライン<sup>\*2</sup> で定義されているアスペクトと、その説明の一部を示す。

WHY: reasons for accident/disaster

WHO AFFECTED: casualties (death, injury), or individuals otherwise negatively affected by the accident/disaster

DAMAGES: damages caused by the accident/disaster

COUNTERMEASURES: countermeasures, rescue efforts, prevention efforts, other reactions to the accident/disaster

我々は、文書クラスタ中のアスペクトを捉えるために最大エントロピー分類器を用いた。最大エントロピー分類器を用いた理由は最大エントロピー分類器の出力は実数値であり、提案手法の枠組みで扱いやすいためである。最大エントロピー分類器の出力は、文  $j$  がどの程度の確率でクラス  $a$  に所属しているかを表しており、我々はこの値を文がどの程度アスペクト  $a$  を反映しているかを表す値とみなす。提案手法では抽出型の要約をおこなうため、文単位でアスペクトの反映度の予測をおこなう。今回用いた素性値を式 (1) に示す：

$$\phi_k(j, y) = \begin{cases} 1, & \text{if } n\text{-gram } k \text{ appears in } j \text{ and } y = a \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

素性値  $\phi_k(j, y)$  は、文  $j$  が  $n\text{-gram } k$  を含み、かつラベル  $y$  が  $a$  であったときに 1 とな

\*1 本稿では便宜上、トピックを TAC によって定義された文書クラスタが属するクラス、カテゴリは Wikipedia の記事が属するクラスとする。

\*2 <http://www.nist.gov/tac/2011/Summarization/Guided-Summ.2011.guidelines.html>

表 1 Wikipedia からトピックに沿った記事集合を集めるために用いたクエリの一部

トピック	クエリ
Accidents and Natural disasters	natural disaster, accident
Attacks	Spree shooting, Massacre, Suicide bombing, Hijacking, Kidnapping
Health and Safety	HIV/AIDS, Organ transplant, Food recall, Product safety, Food borne illness, Health disaster
Endangered Resources	Endangered species, Endangered animals, Water pollution, Environmental issue with forests
Trials and Investigations(Criminal/Legal/Other)	Murder trial, Robbery trial, Trials in, manslaughter, sex crime trial

表 2 本稿で考慮する予め定義されたアスペクト

トピック	アスペクト
Accidents and Natural disasters	WHY, WHO AFFECTED, HOW, COUNTERMEASURES
Attacks	PERPETRATORS, WHY, WHO AFFECTED, HOW, COUNTERMEASURES
Health and Safety	WHO AFFECTED, HOW, WHY, COUNTERMEASURES
Endangered Resources	IMPORTANCE, THREATS, COUNTERMEASURES
Trials and Investigations(Criminal/Legal/Other)	WHO, WHO INV, WHY, CHARGES, PLEAD, SENTENCE

表 3 セクションとアスペクトの対応付けの一部 (トピックは Accidents and Natural disasters)

アスペクト	セクションのタイトル
WHY	Causes, Cause and results, Reasons for crash, Probable cause
WHO AFFECTED	Victims, Fatalities, Injuries
DAMAGES	Damages, Observed damage, Collision
COUNTERMEASURES	Rescue, Emergency response, Recovery

り, それ以外は 0 となる.  $n$ -gram には訓練データ中のユニグラムとバイグラムを用いた, 文  $j$  におけるアスペクト  $a$  の反映度は式 (2) のように表される:

$$p(y|j) = \frac{1}{Z(j)} \exp \left( \sum_k \lambda_k \phi_k(j, y) \right). \quad (2)$$

ただし,  $Z(j)$  は正規化定数で,  $\lambda_k$  は素性  $k$  の重みである.

### 3.2 訓練データのラベリング

藤井ら<sup>6)</sup> は Wikipedia のセクションを利用して, 観点を捉えるための分類器を学習した. 我々も同様に訓練データには Wikipedia を用いた. Wikipedia を訓練データとして使う理由として以下の 3 つの特徴が挙げられる: 1) すべての記事にはカテゴリが付けられていること, 2) 記事はセクションごとに段落が分かれていること, 3) 記事の数が豊富であること. 我々はこれらの特徴を利用して訓練データを作成した. 訓練データ作成から最大エントローピー分類器の学習までの流れを以下に示す.

- (1) 最初に, Wikipedia から文書クラスタのトピックと類似したカテゴリに属している記事集合を生成する. 表 1 に記事集合を収集するためのクエリの例を示す. クエリは文書クラスタ中の各文書を見て人手で決めている. これらのクエリを含むカテゴリに属している記事を Wikipedia から抽出する.
- (2) 次に, アスペクトと対応の取れそうなセクションをラベル付けする. このようなラベル付けをする理由は例えば, セクションタイトル *Causes* はアスペクト *WHY* に関する情報を含んでいると考えられるからである. 表 3 に具体的な対応関係を示す.
- (3) 対応しているセクション中の文集合を訓練データの正例として, アスペクト  $a$  の分類器を学習する. 負例はアスペクト  $a$  以外のアスペクトと対応付けられたセクション中の文集合とした.

このようにセクション単位でラベル付けをおこなった理由はアノテータの負荷を減らすためである. 表 2 に, 本稿で考慮するアスペクトを示す. TAC では, 他にも *WHAT*, *WHEN* や *WHERE* が定義されている. しかし, *WHAT* は事象そのものを表すので頻度や出現位置で十分被覆できると考え考慮しない. さらに *WHEN* や *WHERE* は Wikipedia から訓練データを作成することが難しく提案手法では扱いにくいことを考慮しなかった.

### 3.3 アスペクト被覆のためのモデル化

高村ら<sup>5)</sup>は要約で一般的に重要な要素とされている多様性と文書クラスタがもつ主題との関連性を線形結合したモデル化をおこなった。つまり、要約は冗長な内容を含むべきではないが、要約中に含まれる文は多様性を考慮しすぎて互いに全く別の話題について触れるべきではなく、ある程度似た話題について触れるべきであるという直感に基づいている。提案手法では、高村らのモデルに対してアスペクトの被覆を可能にするための項を線形結合したモデル化を行う。

ここで  $c_i$  を、要約が概念単位  $i$  を含むときに 1、それ以外の場合は 0 となるような変数とする。また  $s_j$  を、要約が文  $j$  を含むときに 1、それ以外の場合は 0 となるような変数にする。さらに、 $o_{ij}$  を文  $j$  が概念単位  $i$  を含むときに 1、それ以外の場合は 0 となるような定数とする。 $w_i$  は概念単位  $i$  が持つ重要度とする。式 (3) の初項、第二項はそれぞれ多様性、文書クラスタが持つ主題との関連性を表している。この二つの項をパラメータ  $\alpha$  で線形結合する。 $\alpha$  は  $[0, 1]$  の実数値で、値が大きいほど多様性を重視し、値が小さいほど主題との関連性を重視するパラメータである。本稿で提案しているアスペクトの被覆をおこなうための項が第三項である。 $z$  という変数は、要約のアスペクトの中で最も低い反映度を表している。要約が持つ任意のアスペクト  $a$  の反映度は、要約に含まれる文集のアスペクト  $a$  の反映度の総和  $\sum_j s_j p_{aj}$  としている。つまりこの変数  $z$  を最大化することで、反映度が最も低いアスペクトのスコアを最大化することになるので、網羅的にアスペクトの被覆をおこなうことをモデル化していることになる。要約の長さ制約  $L$  の範囲内で文を選択して要約を生成する。ただし文  $j$  の長さは  $l_j$  で、ここでは長さ制約  $L$  も各文の長さ  $l_j$  も単語数とする。

最大エントロピー分類器による反映度の予測ではすべての文に実数値が割り当てられるが、文書要約問題を解くときに用いる反映度は各アスペクトについて上位  $N$  件までとする。これは反映度が高いわけではないが、複数のアスペクトである程度の反映度を割り当てられ、結果として不当に要約のアスペクトの反映度を高くしてしまうような文を要約へ含めることを避けるためである。

$$\begin{aligned} & \text{maximize} && (1 - \beta) \left\{ \alpha \sum_i w_i c_i + (1 - \alpha) \sum_j \left( \sum_i w_i o_{ij} \right) s_j \right\} + \beta z, && (3) \\ & \text{subject to} && \\ & && \forall j, \sum_j s_j o_{ij} \geq c_i, \\ & && \forall i, \forall j, s_j o_{ij} \leq c_i, \\ & && \sum_j s_j l_j \leq L, \\ & && \forall a \in \text{aspects}, \sum_j s_j p_{aj} \geq z, \\ & && \forall i, c_i \in \{0, 1\}, \\ & && \forall j, s_j \in \{0, 1\}. \end{aligned}$$

高村らの提案した多様性と主題との関連性の 2 つの項と、アスペクトの被覆を可能にするための項をパラメータ  $\beta$  で線形結合する。 $\beta$  は、 $(0, 1]$  の実数値をとり、値が大きいほどアスペクトの被覆を重視し、値が小さいほどアスペクトの被覆は重視しない。つまり、 $\beta = 0$  のときに、高村らのモデルと一致する。

## 4. 実 験

### 4.1 前 処 理

Wikipedia から配布されている記事の情報が入った xml ファイル<sup>\*3</sup>には、Wiki マークアップのタグが付いているが、WP2TXT<sup>\*4</sup> を用いてタグを除去している。

Wikipedia のテキストを訓練データとして最大エントロピー分類器を学習する際に、訓練データ中の固有表現は不要であるので Illinois Named Entity Tagger<sup>\*5</sup> を用いて固有表現を同定し、固有表現のタグでその固有表現をマスキングしている。以下にマスキングの例を示す。

*Mrs. Morely died while in Lima. → PER died while in LOC.*

また、短い文は有用ではないと考え、文の長さで足切りをして訓練データから単語数が 5

\*3 <http://dumps.wikimedia.org/enwiki/latest/enwiki-latest-pages-articles.xml.bz2>

\*4 <http://wp2txt.rubyforge.org/>

\*5 [http://cogcomp.cs.illinois.edu/page/software\\_view/4](http://cogcomp.cs.illinois.edu/page/software_view/4)

よりも小さい文を除外している。これは、文分割がうまくいかず、文として成立しないような文字列を訓練データとして扱わないようにするためである。

式 (3) において、概念単位は bigram を用いた。また、概念単位の重みには Yih ら<sup>1)</sup> の訓練重みを用いた。この重みは、概念単位が人手で作られた要約に出てくる確率を logistic regression を用いて計算している。logistic regression の訓練データには DUC 2004 のデータセットを利用した。DUC 2004 では、アスペクトは定義されておらず、訓練重みにはアスペクトは考慮されていない。文書クラスタに出てくるすべての語彙は lemmatize をおこなっている。またすべてのアスペクトのスコアの分布は二乗平均が 1 になるように正規化をおこなっている。文書要約問題を解くときに用いた反映度は各アスペクトについて上位 10 件 ( $N=10$ ) までとした。パラメータ  $\alpha$ ,  $\beta$  は交差検定によって決定した。パラメータはトピックごとに決める必要があるため、各文書クラスタに対して、同一のトピックの他の文書クラスタすべてで ROUGE-2 が最も良くなる時のパラメータの組  $\alpha$ ,  $\beta$  を求める。すべてのパラメータの組み合わせで最頻値となるパラメータの組を、そのトピックにおけるパラメータとする。

#### 4.2 データセット

テストデータには TAC 2010 のデータセットと TAC 2011 のデータセットを用いた。TAC 2010 と TAC 2011 における複数文書要約タスクにおいて、参加者にはそれぞれ 44, 46 個の文書クラスタが与えられる。文書クラスタは同じ事象について述べている 10 の新聞記事から構成されていて、100 単語以内で要約を生成することが求められる。また、すべての文書クラスタは 5 つのトピックのいずれか 1 つに必ず属しており、各トピックには複数のアスペクトが予め定義されている。生成される要約はこれらのアスペクトに関する情報を含んでいる必要がある。ただし、文書クラスタがどのトピックに属しているかは参加者に予め知らされている。TAC 2010 と TAC 2011 はタスクに違いがないので二年分のデータを合わせて実験する。文書クラスタのトピックと、文書クラスタの数を表 4 に示す。クラスタ数は二年分の合計となっている。評価には ROUGE-1.5.5 を用いた。検定には対応のある 2 つの母集団間の有意差を検定するためによく用いられる Wilcoxon の符号付き順位検定を用いた。

#### 4.3 実験結果

表 5 に ROUGE-2 による評価値を示す。baseline はアスペクトを考慮しない場合、つまり  $\beta = 0$  の場合である。提案手法を aspect-coverage とする。すべてのトピックにおいてアスペクトを考慮したほうが ROUGE-2 値が改善された。この結果から guided summarization

表 4 トピックごとのテストデータのサイズ

トピック	文書クラスタ数
Accidents and Natural disasters	16
Attacks	16
Health and Safety	22
Endangered Resources	18
Trials and Investigations	18

表 5 ROUGE-2

トピック	baseline	aspect-coverage
Accidents and Natural disasters	0.1102	<u>0.1133</u>
Attacks	0.1307	<u>0.1337</u>
Health and Safety	0.0903	<u>0.0915</u>
Endangered Resources	0.0900	<u>0.0916</u>
Trials and Investigations(Criminal/Legal/Other)	0.1416	<u>0.1469</u>
マイクロ平均	0.1112	<u>0.1140</u>
マクロ平均	0.1126	<u>0.1154</u>

においては、少なくともアスペクトを考慮したほうが質の高い要約を生成できることがわかった。ただし、p 値は 0.059 となり、有意水準 0.1 を満たすが、有意水準 0.05 は満たせなかった。まだ提案手法には改善の余地がある。

## 5. 考 察

図 1 と図 2 にそれぞれ aspect-coverage と baseline による要約結果を示す。これはトピックが *Accidents and Natural disasters* の文書クラスタ D1046H-A の結果である。例えば、“Indonesia’s foreign minister on Friday announced his country will host a January 6 summit of world leaders to discuss the aftermath of Asia’s devastating natural disaster.” という文は、baseline には含まれないが aspect-coverage では要約に含まれる文である。この文が持つ概念単位の重みの総和は 0.316 となっている。この文書クラスタの文の概念単位の重みの総和の平均は 0.406、標準偏差は 0.210 であるのでこの文の概念単位の重みの総和はそれほど大きくない。そのため、文書頻度や出現位置で計算した概念単位の重みの観点からはこの文はあまり重要ではないといえる。しかし、アスペクトを考慮した項によって要約として選択されている。この文はアスペクト *COUNTERMEASURES* を強く反映していると考えられ、そのため ROUGE-2 値も向上したと考えられる。

次に図 3 と図 4 にそれぞれ aspect-coverage と baseline による要約結果を示す。これはト

ピックが *Health and Safety* の文書クラスタ D1044H-A の結果でアスペクトを考慮することで、ROUGE-2 値が baseline よりも下がってしまった例である。この文書クラスタにおいて、ROUGE-2 値が低下してしまった原因として、要約に含まれる概念単位の重要度が高く、アスペクトの被覆を重視できなかったことが考えられる。図 3 中の文の概念単位の重みの総和はそれぞれ、0.858, 3.174, 4.212, 2.975 となっている。この文書クラスタの文の概念単位の重みの総和の平均は 0.7809, 標準偏差は 0.7818 であるところから 4 文中 3 文が高い値であることがわかる。目的関数は、たとえばアスペクトの被覆を考慮するようなパラメータにしても、これらの文が存在するためアスペクトの被覆を考慮できずに ROUGE-2 値が低下してしまったと考えられる。

## 6. おわりに

guided summarization において、アスペクトを明示的に考慮することで ROUGE-2 値が改善されることが確認できた。より良い要約の生成のためには、訓練データをより正確にラベル付けする必要がある。また可読性の向上のためにアスペクトの観点から文の並び替えをおこなうことが有効であると考えられる。

## 参 考 文 献

- 1) Wen-tau Yih, Joshua Goodman, Lucy Vanderwende, and Hisami Suzuki: Multi-Document Summarization by Maximizing Informative Content-Words, in *Proceedings of the 20th International Joint Conference on Artificial Intelligence*, pp.1776–1782 (2007).
- 2) Vasudeva Varma, Praveen Bysani, Kranthi Reddy, Vijay Bharath Reddy, Sudheer Kovelamudi, Srikanth Reddy Vaddepally, Radheshyam Nanduri, Kiran Kumar N, Santhosh Gsk, Prasad Pingali: IIIT Hyderabad in Guided Summarization and Knowledge Base Population, TAC (2010).
- 3) Jaime Carbonell and Jade Goldstein: The use of MMR, diversity-based reranking for reordering documents and producing summaries. In *Proceedings fo the 21st annual international ACM SIGIR conference on Reseach and development in informatino retrieval*, SIGIR '98, pp.335–336, New York, NY, USA, ACM (1998).
- 4) Ryan McDonald: A study of global inference algorithms in multi-document summarization. In *Proceedings of the 29th European conference on IR reseach*, ECIR'07, pp.557–564, Berlin, Hidelberg, Springer-Verlag (2007).
- 5) 高村大也, 奥村学: 最大被覆問題とその変種による文書要約モデル, 人工知能学会論文誌 23 卷 6 号, pp.505–513 (2008).

- 6) 藤井敦, 三條場旭彦: Wikipedia を用いた用語説明のモデル化と事典的検索への応用, 人工知能学会 第 20 回セマンティックとオントロジー研究会資料, SIG-SWO-A803-01, pp.1–8 (2009).

Japan has provided some 500 million US dollar fund for tsunami- affected countries including Indonesia, the Maldives and Sri Lanka. An extremely powerful earthquake and the tsunami that followed devastated many parts of North Sumatra and Aceh on Dec. 26 last year, killing more than 100,000 people there. Uzbekistan on Monday sent a plane with 35 metric tons (39 short tons) of humanitarian aid to tsunami-hit Indonesia, the Foreign Ministry said. Indonesia's foreign minister on Friday announced his country will host a January 6 summit of world leaders to discuss the aftermath of Asia's devastating natural disaster.

図 1 aspect-coveragge による要約 (D1046H-A)

At least 95,000 people were killed by the tsunami in Indonesia. Japan has provided some 500 million US dollar fund for tsunami- affected countries including Indonesia, the Maldives and Sri Lanka. An extremely powerful earthquake and the tsunami that followed devastated many parts of North Sumatra and Aceh on Dec. 26 last year, killing more than 100,000 people there. Uzbekistan on Monday sent a plane with 35 metric tons (39 short tons) of humanitarian aid to tsunami-hit Indonesia, the Foreign Ministry said. The first batch of 100-ton cargoes was airlifted to Colombo, capital of Sri Lanka, Wednesday morning.

図 2 baseline による要約 (D1046H-A)

Those toys were also made in China. About 80 percent of the toys bought in the United States and about 65 percent of Mattel's toys are made in China, according to the Toy Industry Association and the company. In recent months, factories in China have been sources of poisonous pet food sold in stores in the United States, dangerous car tires, and lead paint on the popular Thomas & Friends wooden toys. News that about 1.5 million wooden Thomas & Friends children's toys were being recalled because they were made with lead paint raised some parent s' anxiety levels.

図 3 aspect-coveragge による要約 (D1044H-A)

The recall involves 967,000 plastic preschool toys made by a Chinese vendor and sold in the United States between May and August. "The Consumer Product Safety Commission does not have the resources to catch this type of problem," Mark Pryor, D-Ark. About 80 percent of the toys bought in the United States and about 65 percent of Mattel's toys are made in China, according to the Toy Industry Association and the company. News that about 1.5 million wooden Thomas & Friends children's toys were being recalled because they were made with lead paint raised some parents' anxiety levels.

図 4 baseline による要約 (D1044H-A)