

ChatGPTを用いた抽象化と論理展開による複数コメントの段階的なまとめの生成

Stepwise Summarization of Multiple Comments Using Abstraction and Logical Expansion with ChatGPT

伊東 達希^{1*} 砂山 渡² 服部 峻²
Tatsuki Ito¹ Wataru Sunayama² Shun Hattori²

¹ 滋賀県立大学大学院 工学研究科

¹ Graduate School of Engineering, The University of Shiga Prefecture

² 滋賀県立大学 工学部

² School of Engineering, The University of Shiga Prefecture

Abstract: With the development of the Internet, the need to efficiently extract necessary knowledge from vast amounts of information has increased. However, creating summaries that are appropriately abstracted and incorporate new perspectives requires consideration of the information type and intended use. In this study, we propose a stepwise summary generation system that utilizes abstraction and logical expansion of multiple comments with ChatGPT. Through evaluation experiments, we confirmed that the system effectively supports the creation of summaries that suit the intended purpose.

1 はじめに

近年のインターネットの発達により、誰でも簡単に情報の書き込みや閲覧ができるようになった。例えば、Amazonの商品レビューやYouTubeのコメント、オンラインアンケート結果などがある。

しかし、これらの複数のコメントから新たな知識や結論を得るためには、過度に抽象的なまとめや、具体的なまとめのみに着目すると、新しい知識を得ることは難しくなる。また、新しいアイデアが必要な場合には、論理的な展開を含んだまとめが必要とされる。

こういった課題に対し、本研究ではまとめの抽象化と論理展開に着目する。分析の目的やまとめる人によって異なる抽象度、展開度のまとめが必要とされることが考えられるが、両者に着目した意見集約の研究は見当たらない。異なる抽象度と展開度が必要な例として、新商品の開発をするために、商品のレビューをまとめる場合がある。新商品に取り入れる要素を決める場合には抽象度が低い、具体的なまとめが必要とされる。一方、長期的な会社の方針を決める場合には、抽象度が高い、抽象的なまとめが必要とされる。また、それぞれの意外なアイデアを発想するためには、論理展開を多

く含む展開度の高いまとめが望まれる。そのため、抽象度、展開度別のまとめを作成することは新たな知識を獲得するために有効な手段と考えられる。

そこで本研究では、ChatGPTを利用し、複数コメントの抽象化と論理展開による段階的なまとめ生成システムを提案する。複数コメントの自動集約をChatGPTを利用することで行い、段階的に抽象化、論理展開を行うことで、抽象度別のまとめを生成し評価する。段階的なまとめを生成することでユーザーにとって必要な抽象度、展開度を含んだまとめの一覧を作成し、目的に応じたまとめを生成できるシステムの構築を目指す。

2 関連研究

2.1 複数コメントからの要約生成に関する研究

要約生成の研究は、複数コメントの中から重要と考えられるものを抽出する研究と、複数コメントを抽象化した要約を生成する研究がある。

抽出を行う要約生成の研究として、評価視点別レビュー要約のための重要文候補抽出を行う研究[1]がある。この研究はホテルのレビューから観点ごとの重要文のラベルづけを行い、それを学習したSVMを用いて重要

*連絡先：滋賀県立大学大学院工学研究科 電子システム工学専攻
〒522-8533 滋賀県彦根市八坂町 2500
E-mail: on23tito@ec.usp.ac.jp

文の候補を判定し、抽出している。観点も考慮して集約を行うといった手法は本研究と共通しているが、抽出型の要約は、忠実性が高い反面、論理展開による意外な要約を生成することは難しいと考えられる。そのため、本研究では抽出と抽象化のどちらも取り入れた要約を行う。

抽象化を行う要約生成の研究として、教師なしの要約を生成する研究 [2] がある。この研究は、ホテルのレビューの観点ごとの要約を教師なし学習を用いて生成し、Transformer を用いて生成した擬似要約、レビュー、観点の3つの入力から学習し、要約を生成している。要約生成を行う際に観点をを用いる点では共通しているが、本研究では、大規模言語モデルを利用することで、ホテルのレビューに特化せず、より汎用的なまとめ生成システムを構築する。

2.2 大規模言語モデルによる要約生成に関する研究

近年では大規模言語モデルの登場により、要約に特化したモデルでなくても、プロンプトによる指示のみで複数コメントの集約を行うことができる。GPT-3.5 を使って要約生成を行う研究 [3] がある。この研究は、GPT-3.5 に与える入力が長すぎると要約の精度が下がる問題に対し、レビューのデータセットをチャンクごとに区切ることで要約し、要約の真実性を評価する指標の定義を行った。本研究と ChatGPT を用いて要約を行っている点と、評価指標を定義していることが類似している。しかし、ChatGPT のモデルのアップデートにより、ChatGPT が虚偽の内容を回答する頻度が減っているため、本研究では真実性ではなく、まとめの抽象度に着目した評価を導入する。

ニュース記事の要約において、固有表現の追加により、段階的な抽象度の要約を生成する研究 [4] がある。ここでの固有表現とは、人物名や地名、日付や数値といった情報を指す。この研究は、GPT-4 を用いて、ニュース記事内の主要な固有表現から要約を生成し、要約に含まれていない固有表現を順次追加することで、要約の情報量を段階的に増やすシステムを構築した。本研究と段階的なまとめを生成する点では類似している。しかし、この研究では、固有表現の追加を主な手法としており、要約の情報量は変化するが、要約の観点や構造は一定のままである。本研究では、論理的な展開を含めた抽象化を行うことで、異なる観点でのまとめの生成も行う。

3 複数コメントの抽象化と論理展開による段階的なまとめ生成システム

本章では、複数コメントの抽象化と論理展開による段階的なまとめ生成システムについて述べる。

3.1 本研究における抽象化と論理展開

本節では、本研究における抽象化と論理展開について述べる。複数コメントのまとめを生成する目的は、「主要な知識を求める場合」と「新しい知識を求める場合」の二つに分類される。主要な知識を獲得するために「堅実な解釈」が求められる一方、新しい知識を獲得するためには、「論理の飛躍を含めた解釈」が望まれる [5]。そこで本研究では、ユーザの求める適切なまとめを生成するために、「堅実なまとめ」を抽象化によって生成し、「論理の飛躍を含めたまとめ」は論理展開によって生成する。

本研究では、抽象化を「より広い意味に言い換えること」と「結論に必要な部分を省略すること」と定義する。それを実現する処理として次の二つを用いる。一つ目は「目的に対して重要度が高い1箇所（1単語または1文節）を、より広い意味の言葉に言い換える」。二つ目は、「目的に対して重要度が低い1箇所（1単語または1文節）を省略すること」とする。事前に ChatGPT を用いて行った調査においては、一つ目の処理のみで抽象化を行うと、ChatGPT が入力されたコメントを削除せず、明らかに不要な要素も言い換えることで、過度に細部を維持した抽象化を行う傾向が見られた。そこで二つ目の処理を定義することで、不要な要素が適切に除去されやすくなり、段階的な抽象化を行いやすくなった。また、堅実なまとめの生成を促すために、後述するプロンプトにおいて、1度の抽象化処理につき、1つの処理のみを行うことを指定する。

次に、論理展開の定義について述べる。論理展開は、「元のまとめから連想される新しい視点やアプローチを追加すること」と定義する。ここで新しい視点やアプローチは、元のまとめの直接的な言い換えではなく、元のまとめに関連しつつも、新たな前提や異なる観点に基づく内容とする。抽象化は元のまとめから堅実なまとめを生成するのに対し、論理展開は元のまとめから連想される新しい要素を加えることで意外な発想を促す。例えば、「リンゴが落ちた」を順に抽象化を行うと「果物が落ちた」、「ものが落ちた」、「ものが動いた」となるが、論理展開を行うと、「重力が働いた」といった元のまとめの直接的な言い換えではなく、新たな観点を加えたまとめが生成されることを期待する。

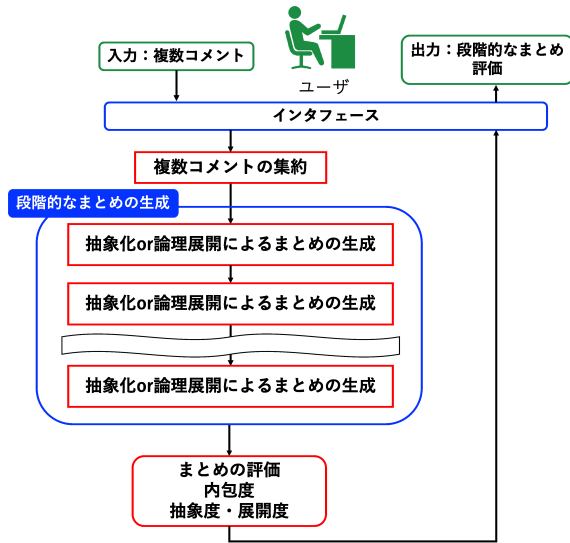


図 1: 複数コメントの抽象化と論理展開による段階的なまとめ生成システムの構成

3.2 システムの構成

図1に複数コメントの抽象化と論理展開による段階的なまとめ生成システムの構成を示す。ユーザーは収集した複数のコメントをシステムのインタフェースに入力する。システムは入力された複数コメントを一つに集約し、その集約結果を抽象化、もしくは論理展開を繰り返すことで段階的なまとめを生成する。さらに、それぞれのまとめの評価を行い、インタフェースを通して段階的なまとめと、それらの評価をユーザーに提示する。本研究での ChatGPT のモデルは、gpt-4o-2024-08-06 を利用する。

3.3 入力：複数コメント

ユーザーが入力する内容は、まとめを行いたい複数コメントと、まとめの目的とする。まとめを行いたいコメントは、1 コメントあたり、20–100 文字程度の意見やレビューとし、コメントの数は2–15 個程度を想定する。これらの想定理由は、複数コメントの集約の際に、文字数やコメント数が多すぎると、重要な要素が抜け落ちてしまう可能性があるためである。本研究では、段階的なまとめ生成に着目しているため、大量のコメントから重要なコメントを抽出する手順はすでに行われているものとする。

また、同じ複数コメントに対しても、観点や目的によって期待する結果が異なるため、まとめの目的を入力する。例えば、「自社商品のレビュー」に対しまとめを作成する場合、「自社商品のレビューから PR 内容を決める」であれば、レビューの中からポジティブな要

表 1: 複数コメントの集約を行うプロンプト（{} 内の文字は変数であり、都度用意された値が代入される）

{目的} という目的において、
{複数コメント} という意見が
指定された文字数になるようにまとめてください。
まとめる方法は、意見の中から共通点を取り出すか、
共通部分を取り出せないもの同士は列挙してください。

入力意見例: りんごを食べたい。トマトを食べたい。
出力意見例: 赤い食べ物を食べたい。

出力は、まとめた意見のみとしてください。
指定する文字数は 100 です。

素に着目する必要がある。一方、「自社商品のレビューから改善策を決定する」であれば、レビューの中からネガティブな要素に着目し、まとめを作成する必要がある。他にも、作成するまとめの方向性の違いや、まとめを行う立場の違いを目的に含むことを想定する。

3.4 複数コメントの集約

本システムでは、入力された複数コメントを ChatGPT を用いて一つに集約する。ここでは、後の処理の段階的な抽象化と論理展開を行うため、可能な限り具体的な情報を保持した集約を行う必要がある。そこで、集約を「意見の中から共通点を取り出すか、共通点を取り出せないもの同士は列挙してください。」と指示することで、入力された複数コメントの要素を可能な限り保持した集約を行う。複数コメントの集約を行うプロンプトを表1に示す。また、具体的な入出力例も含める one-shot プロンプトによる精度の向上も図る。

文字数の指定をしなければ、過度な抽象化が行われ、短い集約結果が返答されることがあったため、複数コメントを集約するのに十分と考えられる文字数として 100 文字を指定する。

3.5 コメント集約結果の抽象化と論理展開による段階的なまとめの生成

本システムでは抽象度と展開度の異なるまとめを生成するために、段階的な抽象化と論理展開を行い、まとめを生成する。また、生成されたまとめが適切あるかを評価するために内包度、抽象度、展開度の3つの指標を提案する。これらの指標を導入することで、ユーザーがインタフェース上で目的にあったまとめを選択しやすくなると考える。

表 2: 抽象化を行うプロンプト（{ } 内の文字は変数であり、都度用意された値が代入される）

{ 目的 } という目的において、{ 集約結果・まとめ } という意見を 1 回だけ抽象化してください。
抽象化では、次の A か B のいずれか片方のみを選択して 1 回だけ実行してください。
A) 目的に対して重要度が低い 1 箇所（1 単語または 1 文節）を省略する。
B) 目的に対して重要度が高い 1 箇所（1 単語または 1 文節）を、より広い意味の言葉に言い換える。
抽象化の結果のみを出力してください。
##抽象化例
美味しいりんごが食べたい。 → りんごが食べたい
出力 りんごが食べたい
文字数は 30 文字を目安にしてください。

3.5.1 抽象化と論理展開の繰り返しによるまとめの生成

前節で得られた集約結果に対し、「抽象化」と「論理展開」の 2 つのプロンプトを繰り返し与え、ユーザの目的に合わせたまとめを生成する。繰り返しの回数はユーザが設定できるが、繰り返しの回数が少なければ、十分な抽象化や論理展開が行われず、回数が多ければ ChatGPT の返答に時間がかかる。これまでの複数回の試行により、処理時間と要約の質のバランスがとれた設定として、デフォルトの組み合わせを「抽象化, 抽象化, 論理展開, 抽象化, 抽象化」の 5 回とする。

「抽象化」は堅実なまとめを行うために、3.1 節で述べた「言い換え」と「省略」の二つの処理として明示的に記述し、1 度の抽象化で一つの処理のみを行う指示をする。抽象化を行うプロンプトを表 2 に示す。具体例として、「おいしいりんごが食べたい」に対して、「省略」の処理を行うと「りんごが食べたい」となる。抽象化の結果のみを出力させることで、抽象化の理由や行った処理の説明といった、意図しない返答を制限する。

一方、「論理展開」は意外なまとめを行うために、元のまとめから連想される新しい視点やアプローチを追加することを指示する。また、論理展開においても、意図しない余計な説明を避けるために、新しい意見のみを出力する指定を行う。論理展開を行うプロンプトを表 3 に示す。例えば、「日常からの脱却」というまとめを展開すると、「非日常の体験」となり、さらに展開すると「魔法が使える」というまとめが得られる。抽象化とは異なり、論理展開を行うプロンプトを用意することで、新たな視点を加えたまとめが生成されることを期待する。

また、それぞれのプロンプトにおいて、one-shot も

表 3: 論理展開を行うプロンプト（{ 目的 } は変数であり、入力された目的が代入される）

{ 目的 } という目的において、与えられる意見から一度だけ連想的な論理展開を行い、新しい意見の一つ答えてください。
論理展開は、元の意見の要点を維持しつつ、新しい視点やアプローチを取り入れてください。
出力は論理展開してできた新しい意見のみとしてください。
具体例
日常からの脱却 -> 非日常の体験
非日常の体験 -> 魔法が使える
文字数は 30 文字を目安にしてください。

しくは few-shot プロンプトを与え、意図しない過度な抽象化や論理展開を防ぎ、より一貫した出力を促す。さらに、まとめの文字数の指定には、Yahoo! ニュース記事の見出しが最大 15.5 文字であることと、BBC ニュースの日本語記事の要約のデータセット [6] の訓練データの平均文字数が約 96 文字であることを参考にし、両者の中間であり、かつ読みやすさを考慮して 30 文字と設定する。

「抽象化」と「論理展開」の組み合わせの割合を変化させた場合のまとめの内容について述べる。繰り返しの回数を固定した場合には、抽象化の割合が高いほど、元のコメント集合の内容を多く保持したまとめが生成される。一方、論理展開の割合が高いほど新しい視点やアプローチが加わるため、元のコメント集合から離れた内容を持つまとめが生成される。これらの組み合わせを自由に設定することで、目的に応じた多様なまとめの生成を行う。

3.5.2 まとめの内包度の評価

本項では前項で生成されたまとめの内包度の評価について述べる。内包度は、システムに入力されたコメントの個数のうち、まとめに使用されたコメントの個数を評価する指標とする。まとめに含まれているかの判断基準は状況や目的によって変化するため、ChatGPT にプロンプトを与え、評価を行った。内包度の評価を行うプロンプトを表 4 に示す。ChatGPT がそれぞれの複数コメントにまとめが含まれるかどうかを「0」と「1」で回答する。初期検証において、リストの長さが入力のコメント数と異なることがあった。その問題に対処するため、総コメント数をプロンプトに含めることで、正しいリストの長さの出力を促す。

内包度 *Entailment* は式 (1) によって定義される。ここで、 n をコメント集合のコメントの総数、 E_k を Chat-

表 4: 内包度の評価を行うプロンプト（{ } 内の文字は変数であり、都度用意された値が代入される）

{ 目的 } というまとめの目的において、
{ 入力された複数コメント } という意見のそれぞれが、
{ まとめ } というまとめに
含まれているかどうかを判断してください。
少しでも含まれている場合は「1」、
含まれていない場合は「0」として、
意見数に合わせて回答してください。
回答には「1」または「0」のみを使用し、
リスト形式で回答してください。
1 行 1 意見とし、意見数は { 総コメント数 } です。
意見数が 6 の時の出力例
[0,1,0,1,1,0]

GPT による返答のリストの各要素とする。

$$Entailment = \frac{\sum_{k=0}^n E_k}{n} \quad (1)$$

内包度の値は $0 \leq Entailment \leq 1$ の範囲を取り、1に近いほど、まとめが入力された多くのコメントを反映していること意味する。一方、0に近い場合は、まとめが複数コメントの内容をほぼ含んでいないことを示す。

ただし、内包度が高ければまとめの質が良いとは限らない。例えば、抽象化を行わず、入力コメントをそのまま羅列したものに対する内包度は1に近づくが、まとめとして適切とはいえずらい。そのため、内包度はまとめの反映率を測る指標として利用するが、他の指標との組み合わせも必要である。

3.5.3 まとめの抽象度と展開度の評価

本研究において、3.1節で抽象化を「より広い意味に言い換えること」と「結論に必要な部分を省略すること」と定義した。これにより、抽象化後のまとめの意味集合は、抽象化前のまとめの意味集合と比較すると拡大すると考えられる。この拡大率を定量的に評価するために、コサイン類似度を用いる。

また、論理展開は、元のまとめの一部を削減しつつ、新たな要素を追加することを含む。この際、論理展開による意味集合の変化は、削減された要素と追加された要素の差分として捉えられる。そのため、論理展開後のまとめと元のまとめとの共通部分の大きさは、類似度として評価できると考えられる。

それらを踏まえて本研究では、抽象度と展開度の定量的評価のために、Embeddingを用いて各まとめをベクトル化し、コサイン類似度を算出する。この手法により、抽象化による意味の拡張度合いや、論理展開による要素の変化を数値的に測定することが可能となる。

Embedding とは、単語や文章を意味ベクトルに変換する手法である。これを行うことで、ベクトル同士のコサイン類似度を求めることができ、単語や文章の類似度を計算できる。本研究では、未知語への対応や、単語ではなく文章をベクトルに変換できること、APIを用いて簡単に利用できるため、OpenAI の Embedding[7] を利用した。このモデルでは、Embeddingを行うことで、文章が 1,536 次元のベクトルに変換される。本システムでは、2025 年 1 月時点で最新モデルの text-embedding-3-large を使用する。

前項で述べた通り、本システムでは、まとめの生成手法を、抽象化と論理展開の2つから選択し、それを繰り返すことで段階的なまとめを生成する。抽象度と展開度の評価においては、まとめ生成手法で抽象化が選択された場合のコサイン類似度が抽象度に積算され、論理展開が選択された場合のコサイン類似度が展開度に積算される。

また、コサイン類似度が高い場合、抽象化による意味集合の拡張度合いや論理展開の変化の度合いは低い。一方で、コサイン類似度が低い場合、直前のまとめからの変化が大きく、抽象化や論理展開の度合いは高い。したがって、コサイン類似度の積を1から差し引くことで、直感的に理解しやすい指標を設計する。

具体的には、各まとめ生成 n において、「抽象化」が行われた場合の抽象化前後のまとめのコサイン類似度を a_n とし、「論理展開」が行われた場合の論理展開前後のまとめのコサイン類似度を b_n とする。ただし、論理展開が選択された場合、そのステップでの a_n の値が未定義となるため、変化なしを意味する $a_n = 1.0$ を代入する。同様に、抽象化が選択された場合の b_n の値も未定義となるため、 $b_n = 1.0$ を代入する。この定義に基づき、全体の抽象度 $Abstractness_N$ を以下の式 (2) で定義する。

$$Abstractness_N = 1 - \prod_{n=1}^N a_n \quad (2)$$

同様に、展開度 $Expansion_N$ も式 (3) を用いて求める。

$$Expansion_N = 1 - \prod_{n=1}^N b_n \quad (3)$$

ただし、複数コメントを集約した結果における a_0 と b_0 は、コサイン類似度を計算する基準が存在しないため、それぞれ 1.0 を代入する。これにより、初期状態では $Abstractness_0 = Expansion_0 = 0.0$ となる。

OptionScreen

まとめの方法を選択してください
3logical

まとめの回数を選択してください
5

詳細設定

設定した抽象化方法: [1, 1, 2, 1, 1]
まとめの目的を入力してください

就活がうまく行った就活生の特徴を知りたい

まとめたい意見を入れてください。

- 自己分析がしっかりできており、将来のビジョンが明確である。
- 「学業に専念していたかどうか。技術力の有無（大学で学習したことが定着しているかどうか）が合否判断を最も左右した。」
- 当社に入社したいという意欲や熱意を面接で強く伝えられる。
- 「短い時間で良く業界・企業研究ができている学生」

まとめてもらう

出力

☆☆☆
自己分析と明確な将来ビジョンを持ち、熱意を伝えられる。学生時代の学びが定義し、業界・企業研究ができている。抽象度：0.0|展開度：0.0

☆☆☆
自己分析と明確なビジョンを持ち、熱意を示せる。抽象度：0.19|展開度：0

☆☆☆
分析と明確なビジョンを持ち、情熱を示せる。抽象度：0.27|展開度：0

☆☆
未来の課題への革新的な解決策を提案。抽象度：0.27|展開度：0.61

図 2: 複数コメントの段階的なまとめ生成インタフェース

3.6 複数コメントの段階的なまとめ生成インタフェース

複数コメントの段階的なまとめ生成インタフェースを図 2 に示す。

画面上部でまとめ生成方法を選択し、中央部で入力を行う。また、画面下部で出力を確認することができる。出力は、一連の処理で生成されたまとめと指標を表示する。指標を全て数字で提示すると視認性が低くなるため、内包度は、4 段階の星の数でユーザに提示する。星の数が多いほど、内包度が高いことを表す。また、集約結果との類似度を視覚的に示すため、まとめのフォントの色の濃淡で提示し、抽象度、展開度はテキストで表示する。

4 複数コメントの抽象化と論理展開による段階的なまとめ生成システムの評価実験

本章では、複数コメントの抽象化と論理展開による段階的なまとめ生成システムが、目的に応じたまとめを得るために有効かを検証した実験について述べる。

4.1 実験方法

本節では、実験方法について述べる。実験で設定した、2つのコメント集合を集約する際の各目的を表 5 に示す。2つのコメント集合と、それらをまとめる 3つの目的に対して、16名の理系大学生、大学院生の被験者に、目的に応じたコメント集合のまとめを、合計 6 回作成してもらう実験を行った。1回の作成の手順は、以下の通りとした。

表 5: 実験で設定した、2つのコメント集合を集約する際の各目的

	動画コメント	商品レビュー
目的 1	動画の特徴を友人に一言で簡潔に伝える際の説明	タオルの特徴を友人に一言で簡潔に伝える際の説明
目的 2	動画を投稿したチャンネル主として、この動画の良さを一言で簡潔に伝えるための説明	タオルの商品開発の視点で、このタオルの良さを一言で簡潔に伝えるための説明
目的 3	動画の特徴を生かした新たな方向性の動画を作る際に、アイデア生成の元になる簡潔なまとめ	タオルの特徴を生かした新たな方向性の日用品を作る際に、アイデア生成の元になる簡潔なまとめ

1. コメント集合をまとめる目的を提示した上で、コメント集合を読んでもらい、そのまとめを作成する。
2. 提案システムによって作成されたコメント集合のまとめを提示し、適切なコメントを選択してもらう。
3. 1. と 2. の内容を踏まえて、再度、コメント集合のまとめを作成してもらう。

実験で用いた目的 1 および目的 2 はサービスを受ける側とサービスを提供する側の観点に分けられる。また、堅実なまとめを必要とする目的 1 と 2 に対し、意外なまとめを必要とする目的 3 を設定した。

コメント集合として用いた動画コメント、および商品レビューは YouTube での再生回数が多い動画、Amazon の売れ筋ランキングで上位の商品を選び、それぞれのコメント集合は評価順で上位 10 件程度抽出した。

手順 2 において、本実験では、抽象度と展開度の指標がユーザの役に立つかを評価するため、内包度の表示は行わず、抽象度と展開度のみを被験者に提示した。また、システムによるまとめ生成の組み合わせをあらかじめこちらで用意することで、システムの利用と同等の結果が得られると考え、まとめの提示、回答は Google forms 上で行った。実験で用意したシステムのまとめ生成手法の組み合わせを表 6 に示す。本システムでは、段階的なまとめを複数回の ChatGPT の応答から得ているが、1 度の応答で段階的なまとめを得ることができれば、システムの待ち時間およびコストを削減できる。そのため、1 回の返答で 1 つの集約と 5 段階の抽象化されたまとめを出力する指示をした ChatGPT による出力も用意した。1 度のリクエストで段階的なまとめを行うプロンプトを表 7 に示す。それぞれの組み合わせにおける 6 個のまとめを 1 セットとして、その

表 6: 実験で用意したシステムのまとめ生成手法の組み合わせ (0 は意見集約, 1 は抽象化, 2 は論理展開, bulk GPT は 1 度の返答で 6 個のまとめを返答する ChatGPT を表す)

組み合わせ
[0, 1, 1, 1, 1, 1]
[0, 2, 1, 1, 1, 1]
[0, 1, 1, 2, 1, 1]
[0, 1, 2, 1, 2, 1]
[bulk GPT]

表 7: 1 度のリクエストで段階的なまとめを行うプロンプト ({ 目的 } は変数であり, 入力された目的が代入される)

{ 目的 } という目的において、与えられる意見を一つの意見に集約し、その集約した意見を少しずつ抽象化することで、抽象度の異なるまとめ意見を 5 個生成してください。集約された意見、抽象化意見 1、抽象化意見 2、抽象化意見 3、抽象化意見 4、抽象化意見 5 の計 6 個の意見を生成してください。

セットをランダムな順番に並べたフォームを作成した。

4.2 実験結果と考察

4.2.1 選択されたまとめの抽象度と展開度からの考察

本項では、選択されたまとめの抽象度と展開度について考察する。手順 2 で選択されたまとめの展開度 (被験者平均) を図 3 に示す。図 3 から、目的「1. 友人に特徴を伝える」において、どちらのコメント集合であっても、抽象度の被験者平均が約 0.08 であった。このことから、「友人に特徴を伝える」という目的に対しては、その特徴をある程度具体的に伝える必要があったと考えられる。

また、目的「2. 制作者として良さを伝える」の商品レビューにおいては、抽象度の被験者平均が 0.18 であった。これはある程度の抽象化を行うことで、商品の良さを漏れなく伝えるためだと考えられる。動画コメントにおいては、被験者平均の値が 0.08 と小さかった。これは、動画の登場人物である、「じいじ」と「ばあば」を用いた抽象化の行われていないまとめが多く選択されることで、抽象度が下がったためであると考えられる。このことから、「制作者として良さを伝える」という目的に対しても具体的なまとめが必要であったと考えられる。

目的「3. 新しいアイデアを生成する」においては、どちらのコメント集合であっても、抽象度の被験者平均が 0.25 以上であった。このことから、「新しいアイデア

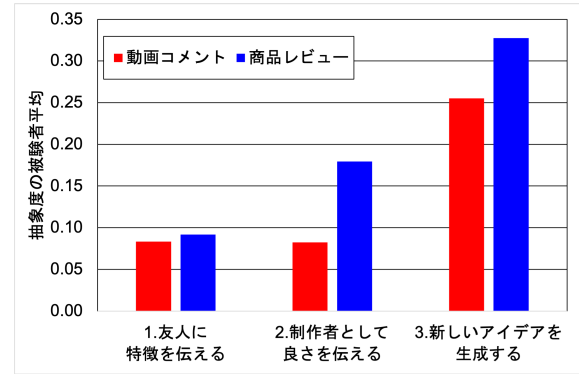


図 3: 手順 2 で選択されたまとめの抽象度 (被験者平均)

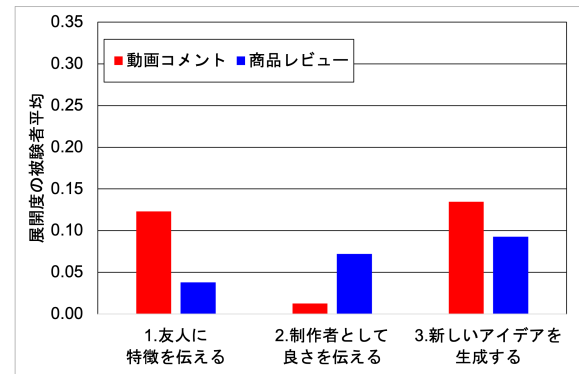


図 4: 手順 2 で選択されたまとめの展開度 (被験者平均)

の生成」に際しては、情報を抽象化する必要性が他の目的と比較して高かったと考えられる。

以上のことから、堅実なまとめを必要とする目的 1 と 2 に対し、意外なまとめを必要とする目的 3 とでは要求される抽象度が異なることを確認した。よってコメント集合や目的が異なれば、異なる抽象度のまとめが必要となる可能性があり、様々な抽象度のまとめを提示することには意味があると考えられる。

同様に、手順 2 で選択されたまとめの展開度 (被験者平均) を図 4 に示す。図 4 より、どの目的においても展開度の被験者平均はおおよそ 0.1 前後であり、顕著な差は見られなかった。このことから、本実験においては、被験者は論理展開を加えたまとめよりも、堅実なまとめを選ぶ傾向があったと考えられる。今回の評価実験では「一言で簡潔に伝えるための説明」といった指示であり、具体的な文字数の指定がなかったことに加え、複数選択可であった。それにより、多くの文字数でより多くの情報を持つまとめ (低抽象度, 低展開度) が選択される傾向が強まったと考えられる。

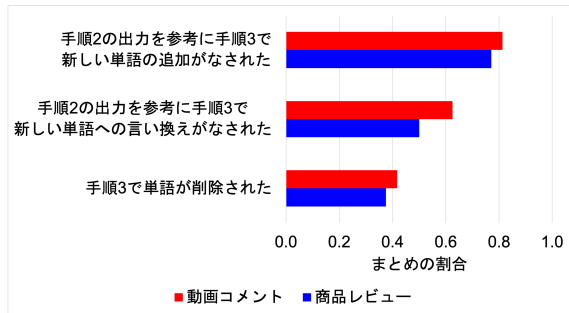


図 5: 手順 1 から手順 3 におけるまとめの変化の重複あり分類の結果

4.2.2 まとめの提示により作成されたまとめの変化についての考察

この項では、まとめの提示により作成されたまとめの変化について考察する。前節の手順 1 と手順 3 において、被験者ごとのまとめの変化を「新しい内容の単語を追加」「内容を維持したまま、新しい単語を用いて言い換え」「単語の削除」の 3 つに重複を許して分類した。そのコメント集合ごとの割合を図 5 に示す。ここから、約 8 割のまとめにおいてシステムによるまとめの内容を用いた単語の追加が行われ、およそ 6 割のまとめが内容を維持したまま、単語の言い換えを行った。このことから、被験者はシステムによるまとめを参考にしつつ、自らの表現に適した単語を追加、修正することができたと考えられる。

また、約 4 割のまとめにおいては、単語の削除が行われた。これは、システムによる抽象化されたまとめを参考に、目的に対して不要な内容を整理できたためであると考えられる。特に、手順 1 でのまとめの文字数が多い被験者においてこういった変化が多く見られた。

これらを踏まえ、システムによる段階的なまとめの提示は、まとめる目的に対して適切な内容を追加し、不必要な内容を削除しながら、ユーザにとって適切な表現へと修正する支援を行える可能性があると考えられる。したがって、本システムはまとめる目的に応じた柔軟なまとめ作成を支援する方法の一つとなりうる。

5 おわりに

目的に合わせた段階的なまとめを生成するために、複数コメントの抽象化と論理展開による段階的なまとめ生成システムを提案した。抽象化と論理展開を定義し、ChatGPT を用いた段階的なまとめ生成によって堅実なまとめと意外なまとめを生成し、それらを評価するシステムを提案した。

また、段階的なまとめの提示がユーザにとって目的にあったまとめの作成に有効であるかを確認する評価

実験を行った。その結果、目的ごとに必要とされる抽象度が異なることがわかった。また、まとめの提示によりユーザにとって必要な要素を追加し、ユーザにとって適切な表現へ言い換える支援として機能する可能性を示せた。

参考文献

- [1] 小池惇爾, 松吉俊, 福本文代: 評価視点別レビュー要約のための重要文候補抽出, 第 18 回言語処理学会年次大会, pp.1188–1191, (2012)
- [2] Amplayo, Reinald Kim, Angelidis, Stefanos, Lapata, Mirella: Aspect-Controllable Opinion Summarization, Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP 2021), pp.6578–6593, (2021)
- [3] Bhaskar, Adithya, Alex Fabbri, and Greg Durrett: Prompted opinion summarization with GPT-3.5, Findings of the Association for Computational Linguistics (ACL 2023), pp.9282–9300, (2023)
- [4] Griffin Adams, Alex Fabbri, Faisal Ladhak, Eric Lehman, and Noémie Elhadad: From Sparse to Dense: GPT-4 Summarization with Chain of Density Prompting, In Proceedings of the 4th New Frontiers in Summarization Workshop, pp.68–74, (2023)
- [5] 砂山渡: フリーソフト TETDM で学ぶ実践データ分析, コロナ社, (2020)
- [6] Mandy Guo, Zihang Dai, Denny Vrandečić, and Rami Al-Rfou: Wiki-40B: Multilingual Language Model Dataset, In Proceedings of the Twelfth Language Resources and Evaluation Conference, pp.2440–2452, (2020)
- [7] OpenAI Platform Vector embeddings: ([URL](https://platform.openai.com/docs/guides/embeddings/)) <https://platform.openai.com/docs/guides/embeddings/> (2025 年 1 月 23 確認)