

仕様書パラグラフ構成文の論理構造同定に基づく 試験ケース・試験パラメータの導出支援

緒方佑亮¹ 久代紀之¹

概要： システム試験における試験設計は、仕様書に記載された各文をテスト設計者が熟読し、“試験条件⇒期待動作”対を試験ケースとして抽出することで達成される。
仕様書には、自然言語特有の曖昧さや論理的な誤りを含んでおり、その修正・補完に膨大な工数が必要であるととも、試験ケース設計作業が手作業で行われるため、テスト技術者のスキルによる品質のばらつきが大きかった。
本研究では、自然言語処理技術を用い仕様書記載文から“条件⇒期待値”包含関係を含む文をキーセンテンスとして自動抽出する。さらに、抽出したキーセンテンスを中心として、パラグラフ構成文を論理的に構造化し、その論理的構造に基づき、試験ケースの試験条件・期待動作パラメータの導出の支援を行うことで、試験ケース設計の効率化・高品質化を試行する。

キーワード： 文間構造、談話処理、試験ケース

Support for derivation of test cases and test parameters based on identification of logical structure of specification paragraphs

YUSUKE OGATA^{†1} NORIYUKI KUSHIRO^{†1}

Abstract: Test design in system testing is accomplished by the test designer carefully reading each sentence in the specification and extracting the "test condition ⇒ expected behavior" pairs as test cases.
The specifications contain ambiguities and logical errors characteristic of natural language, which require an enormous amount of man-hours to correct and supplement, and since the test case design work is done manually, there is a large variation in quality depending on the skills of the test engineers.
In this research, we use natural language processing technology to automatically extract sentences that contain the "condition ⇒ expected value" inclusion relation as key sentences from the specification text. Furthermore, we try to improve the efficiency and quality of test case design by logically structuring the paragraph sentences around the extracted key sentences and supporting the derivation of test conditions and expected behavior parameters for test cases based on the logical structure.

Keywords: Discourse relation, Discourse analysis, Test design

1. はじめに

システムを開発する際に、システムの最終品質を担保するためには、そのシステムに対してシステム試験を実施する。システム試験を実施するためには、テスト仕様書が必要となるが、テスト仕様書作成は、テスト技術者が、自然言語で記載されたシステム仕様書を熟読し、“試験条件⇒期待動作”対を試験ケースとして抽出することで達成される。一方、システム仕様書には、一般的に自然言語特有の曖昧さや論理的な誤りを含んでおり、その修正・補完には膨大な工数が必要である。さらに、試験ケース設計作業は手作業で行われるため、テスト技術者の経験やスキルにより品質のばらつきという課題が存在していた。

先行研究[1]では、これらの課題を解決するために、自然言語で記載されている文を形式的な記述に変換し、試験ケースを自動生成するための研究を実施した。しかし、形式

記述への変換プロセスが単文ごとに行われているため、仕様書に記述される文章の文間の関係性が把握できなくなってしまう。試験ケース作成に必要な“試験条件⇒期待動作”対を抜き出すことができたとしても、試験条件・期待動作に対する前提や制約を記述した試験ケースに対する補足説明の内容を把握できないといった課題が存在した。

そこで本研究では、“試験条件⇒期待動作”対の含意関係を中心としてパラグラフを構成する各文の論理関係を文間の類似性により算出し、その構造を可視化することで、その論理関係を構造化・可視化した。加え、構造化した結果を基に、試験ケース設計において考慮すべきパラメータの導出支援を行うことで、前述の課題の解決を試みた。

2. アプローチ

本研究における仕様書パラグラフの構成各文の論理構

¹ 九州工業大学
Kyushu Institute of Technology.

造を同定し、試験パラメータの抽出支援を行うための提案プロセスを図1に示す。以降図中の赤文で示すプロセスについて、説明する。

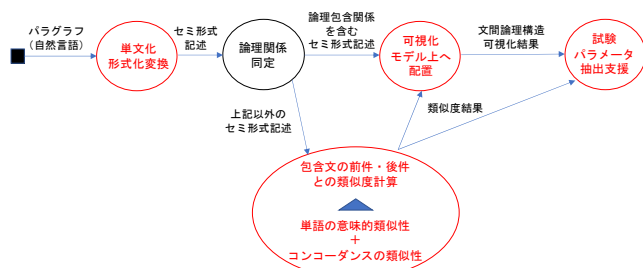


図1 提案プロセス

2.1 単文化・形式化変換

文章から欠落情報の抽出を行うために、自然言語で記載されている文を「関係語（主体，対象，制約）」の形式の命題プリミティブと呼ぶ単位で表現するセミ形式記述[1]に変換する。セミ形式記述に変換した際、文の中で主体や対象の表記漏れが存在する場合には、該当する漏れの箇所を「??」で表現する。これにより、レビュー時に、記述漏れの発見および漏れの補完を促すことができる。

例えば、「設定温度を100度に設定する。」という文をセミ形式記述化すると、「設定する(??,設定温度,100度)」と形式的に表現される。この際に、「??」に該当する箇所が表記漏れであることを示しており、主体に相当する語あるいは句が欠落していることが分かる。このような欠落情報をセミ形式記述変換後のレビューにて、レビュワーに対し補完を促すことができ、元の文章に含まれていた曖昧性を排除することが可能となる。

また命題プリミティブ同士は論理演算子(Not「!」, And「&」, Or「|」, Imply「->」)でお互いを接続することで厳密な関係性を表記することが可能となっている。

本研究において構造化する際の基準としている「Imply」関係（包含部分）の同定は、実システム仕様書の分析結果から抽出した含意関係のルールに基づき判定している。具体的なルールの代表例としていくつか挙げると、形態素解析後の得られる文節の末尾が「場合」、「時」、「中」で終わる場合等がある。

2.2 概念類似度算出

2.2.1 日本語 WordNet

類似度の概念算出には、日本語 WordNet を適用した。

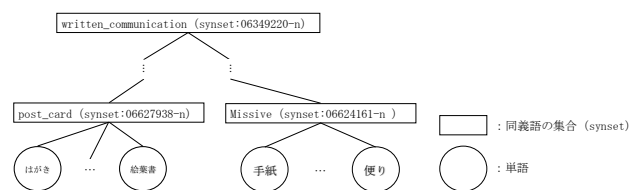


図2 日本語 WordNet の構造例

日本語 WordNet[2]の構造は、単語を類義関係の集合としてグループ化を行い、この集合を synset と呼び基本単位と

している（図2）。

単語の類義関係の集合（以下概念と呼ぶ）に対しては、概念間の関連がリンク付けされており、意味的な上位語

(Hypernym) や下位語 (Hyponym) などが定義されている。概念にはそれぞれ分類番号が割り当てられており、「はがき」の属する概念「post_card」には「06627938-n」が割り当てられている。分類番号の末尾のアルファベットは日本語 WordNet 上で名詞であれば「n」、動詞であれば「v」として登録されている。つまり、品詞毎に、独立した木構造のネットワークを構成している。

それぞれの品詞ごとに関連がリンク付けされている。

この概念構造を活用し、(1) 式で概念類似度を算出する。

$$sim(w_1, w_2) = \frac{2 * S_c}{S_{w_1} + S_{w_2}} \quad (1)$$

ここで単語 w_1 , w_2 において概念類似度を算出する場合、対象とする単語の属する概念の深さ（根を深さ1とする）は S_{w_1} , S_{w_2} であり、 S_{w_1} , S_{w_2} の共通とする上位ノードの深さは S_c である。

2.2.2 2つの類似性

本研究では、2つの観点から類似度を算出し、足し合わせることで、文と文の関連性を導出する。

2.2.2.1 単語の意味的類似性

単語の意味的類似性としては例えば、文脈において「はがき」と「手紙」という単語は人間が理解する際には似たような意味、あるいは同じ内容を意味している単語であると認識できる。これは図2の構造例を参照すると、「はがき」や「手紙」が属するそれぞれの概念にはその上位概念に、共通する概念が存在している。この類似性により多少のニュアンスの違いや書き方の違いはあるものの、意味の近似性があるため、「はがき」と「手紙」においては、意味的な類似性が高いと判断される。文を構成する単語毎に意味的な類似性を算出することで、単語間の類似性をもとに、それら語群で構成される文と文の類似性を決定していく。

2.2.2.2 コンテキストの類似性

コンテキストの類似性では、ある語に対してよく共起する語に関する類似性を考慮する。これは自然言語処理の分布類似度の考えを借用した。分布類似度は、「文脈の似ている語は類似している」ということに着目し、「共起する語が似ていれば類似している」という考え方に基づいた尺度である。例えば「はがきを出しに行きました。」という文と「郵便局に行きました。」の2文は、コンテキスト的には、似たようなことを意味していると解釈可能である[3]。これは、「はがき」と「郵便局」は、共起する可能性が高く、それにより共起する語で構成される文脈も意味としても類似している。あるいは強い関連性があると判断する。共起する単語の取得方法として、概念毎に、WordNet が定義する概念定義文を構成する単語を活用する。概念定義文はその概

念に属している単語の共通点を抽象化し、一般的な意味として言い表した言葉であることから、共起する単語の収集や共起しやすい単語の概念の発見にも繋がると考えた。

2.2.3 類似度の算出方法

2.2.3.1. 概念 ID の取得と概念類似度

まず、記載されている単語が日本語 WordNet において、どの概念に属しているかを取得する。図 3 の入力されたセミ形式記述例をもとに説明していく。{ex.1} の後件の「呼鈴」という単語を日本語 WordNet で検索した結果が図 4 である。

- (1) $w_1(w_2, w_3)$. {ex.1} $\frac{\text{押す(ユーザー, 電源ボタン)}}{P_1} \& \frac{\text{鳴らす(機器, 呼鈴)}}{P_2}$
- (2) $w_i(w_{ii}, w_{iii})$. {ex.2} $\frac{\text{鳴る(ベル, ??, 3秒間)}}{P_3}$

図 3 セミ形式記述の基本構造 (左) と例文 (右)

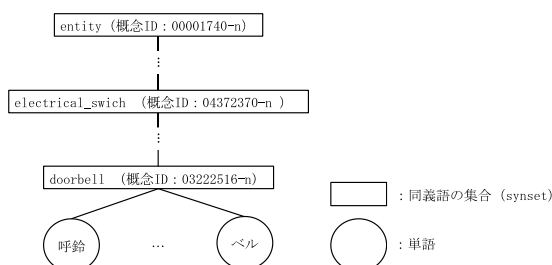


図 4 日本語 WordNet における呼鈴の定義

図 4 より、「呼鈴」は「doorbell」という概念に属しており、その概念に割り当てられる ID が「03222516-n」ということを取得する。「ベル」に関しても「呼鈴」と同じ「doorbell」という概念に属しており、概念 ID としては「03222516-n」が取得できる。

次に、概念類似度を算出するためには、属する概念の深さと 2 つの概念間での共通する上位概念の深さを求める。まず、属する概念の深さ (S_{w_1} , S_{w_2}) に関しては、「呼鈴」と「ベル」の場合、最上位概念を深さ 1 とすると、最上位概念から「entity, ..., electrical_switch, ..., doorbell」という一連の関連をたどり、深さ 12 として定義がなされている。共通概念の深さ (S_c) は、「doorbell」という概念が共通であり、その深さも同様に 12 である。

これらの情報と (1) 式より「呼鈴」と「ベル」の単語の意味的類似性を算出すると、

$$\text{sim}(\text{呼鈴}, \text{ベル}) = \frac{2 * 12}{12 + 12} = 1$$

として、単語間の類似度が得られる。

一方で、「呼鈴 (概念 ID: 03222516-n)」と「鳴る (概念 ID: 02770535-v)」のように名詞と動詞の組み合わせの場合、品詞が違うことから日本語 WordNet の構造上では、品詞毎に独立した木構造で概念を管理しているため、共通する概念を見出すことができず、類似度は 0 となってしまう。

上記作業を、命題プリミティブを構成する単語である

$w_1 \sim w_3$ と $w_i \sim w_{iii}$ の組み合わせにおいて同様の処理を行う。日本語 WordNet において、定義されていない単語に関しては、概念 ID を取得できないため、定義されていない単語として後に文字列の一致で類似度を加算する。

2.2.3.2. 共起する単語の取得と類似度

共起する単語の取得を次に行っていく。2.3.3.1 節において単語が属する概念が取得できているものとする。本ステップでは、その概念に存在する概念定義から共起する単語を取得する。「呼鈴」の属する「doorbell (03222516-n)」という概念には、「押した際に、音が鳴る、またはブザーが鳴って合図するドアの外側についた押しボタン」という概念定義が存在する。概念定義を形態素解析機 JUMAN[4]で名詞と動詞の抽出を行い、得られた単語を共起する単語として取得する。「呼鈴」の場合は、「押す、音、鳴る、ブザー、合図、する、ドア、外側、つく、押し、ボタン」を共起する単語として取得する。

同様の処理を、「鳴る」で実施すると、「鳴る」が属する概念は「thunder (02770535-v), sound (02179518-v), peal (02180898-v), sound (02180529-v)」の 4 つであり、定義文としては「雷が聞こえている例である、鳴り響かせる、大きく朗々と鳴り響く、音によって知らせる」が定義されている。

これら定義文から名詞と動詞の抽出を行った「雷、聞こえる、例、鳴り響く、音、よる、知らせる」を共起語 (名詞・動詞) として取得する。

しかしこの際、以下の問題点が存在する。

1. 定義文に出現する単語において「する」や「人」など多くの定義文で出現する単語を含めて共起する単語として取得しようと、関連性の低い単語間であっても類似性や関連性が存在していると算出される。
2. 多義性である単語では、複数の概念に属することから、共起語の数が極めて大きくなってしまう。

そこで上記 2 つの問題を解決するために、まず 1 に関しては、多くの定義文で出現する単語以外から選定するために TF-IDF 値を活用し、高頻出な単語と、ある単語に対して特有の単語を明瞭化し先の問題の解決を図る。TF-IDF の算出には、単語とその単語とよく共起する単語のタプルが必要となるが、先行研究[5]で調査して得られたワードベクトルデータを活用し、4,385 語に対して出現する単語 35,258 語 (重複を含む) を基にして単語の頻出度を算出した。

次に 2 に関しては、先程算出した単語の頻出度のデータと、共起する単語数から以下の 4 ステップで選定数を決定し、単語を選定した。

1. 単語の頻出度より共起する単語を降順でソートする。
2. 共起する単語を日本語 WordNet 上での名詞、動詞のリストにそれぞれ振り分ける。
3. 共起する単語の総数に応じて抽出数を決定し、名詞、動詞のリストそれぞれから抽出する

総数が 10 未満：リストから 1 つずつ抽出

総数が 10 以上：リストから (総数/10 + 1) つずつ抽出 (小数点以下切り捨て)

4. 名詞、動詞の単語リストを共起する単語とする。

上記ステップを「呼鈴」に対して実施した結果は図 5 のようになる。選定する前の共起する単語は「押す、音、鳴る、…、つく、押し、ボタン」と 11 個の単語が存在していたが、名詞と動詞からそれぞれ 2 つずつ抽出した、「ブザー、合図、鳴る、押し」の計 4 個の単語を共起する単語として扱う。また、「する」などの頻出度の高い単語については、除外できていることがわかる。

“呼鈴”：{
“押す”：0.3238898501504031,
“音”：0.25718693100538215,
“鳴る”：0.3786225766347634,
“ブザー”：0.3786225766347634,
“合図”：0.3512562133925832,
“する”：0.1291651159770322,
“ドア”：0.3512562133925832,
“外側”：0.2904374151236218,
“つく”：0.26237224927541525,
“押し”：0.3672645096703725,
“ボタン”：0.3672645096703725
}

→

“呼鈴”：{
“ブザー”：0.3786225766347634,
“合図”：0.3512562133925832,
“鳴る”：0.3786225766347634,
“押し”：0.3672645096703725
}

図 5 共起する単語の選定例

同様の処理を、「鳴る」で実施すると、選定する前の共起する単語は「雷、聞こえる、例、鳴り響く、音、よる、知らせる」と 8 つの単語が存在していたが、名詞と動詞からそれぞれ 1 つずつ抽出した、「雷、鳴り響く」の計 2 個の単語を共起する単語として扱う。

ここまでの処理により共起する単語として、「呼鈴」に関しては「ブザー、合図、鳴る、押し」を選定し、「鳴る」に関しては「雷、鳴り響く」を選定した。次に得られた共起単語を使いコンコーダンスの類似性を求めていく。共起する単語として得られた単語の組み合わせに対し、それぞれ 2.3.3.2 節で説明した方法で単語間の類似度を算出する。「呼鈴」と「鳴る」の例で説明すると、「ブザーと雷」、「ブザーと鳴り響く」、「合図と雷」…「押し、鳴り響く」と計 8 つの組み合わせで単語間の類似度を求める。得られた結果のうち最も類似度が高い値を「呼鈴」と「鳴る」のコンコーダンスの類似度として扱う。共起する単語での組み合わせの中で、共通する上位概念の深さが 2 である「鳴る (深さ: 3) と鳴り響く (深さ: 3)」の類似性が最も高くその値は (1) 式より 0.666 となる。

2.2.3.3. プリミティブ間の類似性の算出

2.3.3.1 節と 2.3.3.2 節の結果を活用することで、単語の意味的類似性とコンテキストの類似性を算出した。これら 2 つの類似性を合算した値をプリミティブ間の類似性として採用する。例えば「呼鈴」と「鳴る」の単語の類似度は、意味的類似性の類似度 0 と、コンコーダンスの類似度 0.666 を足した値 0.666 であると求まる。

プリミティブ間の類似性は、セミ形式記述を構成する単語ごとに算出された単語の類似性を合算することで算出する。図 3 の {ex.1} {ex.2} を例にして、2 つの文の文間の類似性

を求める。まず「 P_1 と P_3 」について考える。「押す」「ブザー」「電源」それぞれに対して {ex.2} の「鳴る、ベル、3 秒間」のうちもっとも単語の類似性が高い値を算出する。また、{ex.1} に関しては論理積で結ばれた命題プリミティブが存在している。このように論理積や論理和で結合されている内容に関しても同様の処理を行う。「 P_2 と P_3 」について考えると、「鳴らす」「機器」「呼鈴」それぞれに対して {ex.2} の「鳴る、ベル、3 秒間」のうちもっとも単語の類似性が高い値を算出する。その結果、命題プリミティブ間の類似性が算出できるため、「 P_1 と P_3 」「 P_2 と P_3 」の類似度を足し合わせた結果を文間の類似性として算出し活用する。

2.2.4 パラグラフの構造化

2.2.4.1. 可視化モデル

2.3.2 節で得られる文間の類似度を用いて、文間の構造を可視化するに当たり、Toulmin's Model を参照し作成した可視化モデルを図 6 に示す。Toulmin's Model とは、一貫性のある議論を組み立てるための枠組みを提供し、議論における情報抽出漏れの検知や構造的かつ視覚的に関係性を可視化することのできる議論フレームワーク[6]である。

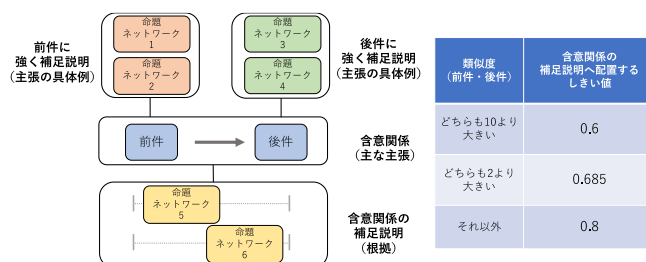


図 6 可視化モデルと配置ルール

含意関係を中心として、「前件に強く補足説明している内容、後件に強く補足説明している内容、含意関係自体に補足説明されている内容」の 3 つの構造に分類することで可視化した際に、直感的に文の論理構造を把握することができる。分類し配置する際のルールとしては、2.3.3 節で算出された前件と後件に対するそれぞれの類似度において、高い値に対する低い値の比率により決定している。例えば図 6 における「命題ネットワーク 5」を例に説明する。「命題ネットワーク 5 と前件」「命題ネットワーク 5 と後件」の類似度がそれぞれ 5.0 と 3.5 の場合、 $5.0 : 3.5 = 1.0 : 0.7$ となり図 6 の配置ルール中段より、「含意関係の補足説明」として配置する。配置ルールに満たない比率である場合には、類似度の高く算出される方へ配置される。

可視化する際には、セミ形式記述を可視化し論理関係の把握支援をするために作成された、命題ネットワーク[1]を活用した。次にその命題ネットワークについて説明する。

2.2.4.2. 命題ネットワーク

セミ形式記述はテキストベースで記述するものであるが、文内における構造が複雑化した際には、構造の把握が難しくなる。そこで命題プリミティブ間の論理関係を図式化す

るために命題ネットワークを用いて可視化を行う。図7に命題ネットワークでの表現方法について示す。セミ形式記述における「関係語」を中心として、左側に「主体」、右側に「対象」、下側に「制約」を配置して可視化を行う。

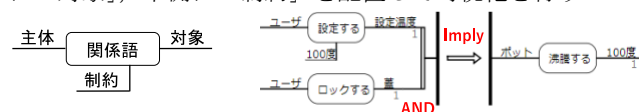


図7 命題ネットワークの構造とその例

2.2.5 試験パラメータ抽出支援

これまでの過程において、含意関係と他の文の文間構造を同定することで、文単位で関連性の抽出を行うことができた。しかし試験パラメータでは、文の中に記載されている「どのような状態値を取りうるか」や「状態の具体的な継続時間」といったより細かい粒度での情報が必要となる。そこでセミ形式記述変換で得ることのできる単文単位での構造に着目し、文の中のどのような要素がより含意関係に対して、関連性を持った情報であるかを算出する。そのために2.2.3節で算出された類似度と、2.3.4節で得られる文間構造同定結果を基にして試験パラメータの抽出支援を行っていく。

文間構造を同定したことで、次の4つの組み合わせから含意関係に対してより関連の存在する単文を見つける。

- A) 「前件」と「前件に強く補足説明」
- B) 「前件」と「含意関係の補足説明」
- C) 「後件」と「後件に強く補足説明」
- D) 「後件」と「含意関係の補足説明」

4つの組み合わせごとに、得られた類似度を基にして前件あるいは後件に対して類似性の高いものから順に命題プリミティブをcsvファイルで出力する。

3. 検証

3.1 題材

今回検証をする題材には、話題沸騰ポット（GOMA-1015型）要求仕様書第7版[7]（以下沸騰ポット仕様書）を用いた。沸騰ポット仕様書のうち、第2章の操作要求仕様に記載されている内容をパラグラフとして見立て入力とした。例えば、沸騰ポット仕様書のpot-220には図8のような記述がある。

2. 2 章	要求	pot-220	アイドル状態で、蓋を閉じたら、水位を確認し、条件に合えば沸騰行為をする。
	理由		蓋を閉めるという行為で加熱（沸騰）の指示をしたい。
	説明		沸騰行為の詳細は、3章の「温度制御行為」に記載する。 蓋センサーがonになって3sec経過するのを待つ理由は、注水やポットの移動の直後に、水面が波打っている状況が考えられるので、水面状態が安定する時間を想定したためである。
	<蓋「閉」を確認する>	pot-220-1	蓋センサーが3sec以上onになったら、蓋が閉じられたと判断する。
	<水量適正時の処理>	pot-220-2	蓋が閉じられ、水量が適正な場合、沸騰行為をする。 【説明】水量については、pot-280を参照。
	<水量異常時の処理>	pot-220-3	蓋が閉じられても、水量が異常な場合、状態はアイドルのままである。

- [1] 蓋を閉めるという行為で沸騰の指示をするために、アイドル状態のときに、蓋を閉じたならば、水位を確認して条件下で沸騰行為をする。
- [2] 沸騰行為の詳細は、3章の「温度制御行為」に記載する。
- [3] 蓋センサーがonになって3sec経過するのを待つ理由は、注水やポットの移動の直後に、水面が波打っている状況が考えられるので、水面状態が安定する時間を想定したためである。
- [4] 蓋センサーが3sec以上onになったら、蓋が閉じられたと判断する。
- [5] 蓋が閉じられ、水量が適正な場合、沸騰行為をする。
- [6] 蓋が閉じられても、水量が異常な場合、状態はアイドルのままである。

図8 話題沸騰ポット仕様書例と文章化した結果

表の中で「要求」と記載されている部分が沸騰ポットの機能として満たすべき内容であり、試験ケースとして検証すべき内容であると考え、この要求文に含意関係を含んでいる機能を構造化とパラメータ抽出支援の題材とした。また、表内の「理由」の扱いとして、表を用いているため「要求」と分けて記載されているが、本来であれば要求内に記載可能な内容であるため、「要求」と「理由」に関しては、1文で記載し直した。図8の例では、「要求」には「アイドル状態で、蓋を閉じたら、水位を確認し、条件に合えば沸騰行為をする。」記載され、「理由」には「蓋を閉めるという行為で加熱（沸騰）の指示をしたい。」と記載されている。これらの2文を本研究では、「蓋を閉めるという行為で沸騰の指示をするために、アイドル状態のときに、蓋を閉じたならば、水位を確認して条件下で沸騰行為をする。」と1文にまとめた。その結果pot-220の場合では6文で構成されたパラグラフとして扱い提案のプロセスにおける入力とした。この作業のもと検証の題材として選定された8つの機能についてそれぞれ次節の項目を評価した。

3.2 評価項目

本研究では、仕様書のパラグラフ構成文の論理構造同定により得られる構造化した可視化結果と、試験パラメータとして考えられる内容の選定を実施しているため、上記2点に関してそれぞれ再現率を算出した。まず、論理構造同定における評価としては、算出された類似度により可視化モデルへの配置を行った結果に対して人手での配置と比較し、人間の解釈による構造化に対するアルゴリズムによる配置結果の再現率を求めた。人手による配置に関してpot-220を例に説明する。pot-220のパラグラフは図8において得られた文章である。人手での配置の際に考慮した内容を説明するためにわかりやすく文の構成を色分けした結果と人手での配置結果が図9である。セミ形式記述へ変換した際に得られる含意関係において、1文目の赤色が前件であり、後件が青色で示している。その中で共通する単語を2文目から6文目においても色付けを行い、構造化する際この情報をもとにして配置を行った。例えば、4文目に関しては、

「蓋」という単語で赤色のみが着色されているため、前件に強く関連性があると判断し、「前件に強く補足説明」へ配置する。5文目に関しては、赤色も青色も混同している状態であり、内容としては含意関係自体に関する内容であるため「含意関係の補足説明」へ配置するといった形で人手での配置を実施した。

- [1] 蓋を閉めるという行為で沸騰の指示をするために、アイドル状態のときに、蓋を閉じたならば、水位を確認して条件下で沸騰行為をする。
[2] 沸騰行為の詳細は、3章の「温度制御行為」に記載する。
[3] 蓋センサーがonになって3sec経過するのを待つ理由は、注水やポットの移動の直後に、水面が波打っている状況が考えられるので、水面状態が安定する時間を想定したためである。
[4] 蓋センサーが3sec以上onとなったら、蓋が閉じられたと判断する。
[5] 蓋が閉じられ、水量が適正な場合、沸騰行為をする。
[6] 蓋が閉じられても、水量が異常な場合、状態はアイドルのままである。

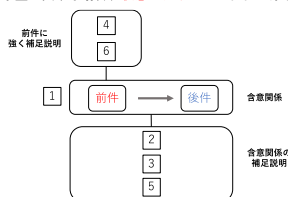


図9 pot-220の内容(上)と人手による配置(下)

また、パラメータ抽出に関しても人手との比較を行うが、人手での試験パラメータの候補として抽出する指針としては、以下の指針を設定した。

- 「～状態」など変数として考えることのできる内容とその変数に代入されうる項目
- 前件(試験条件)、後件(期待動作)と同時あるいは事前、事後に発生するイベント
- 前件(試験条件)、後件(期待動作)において成立して置かなければならない条件

これらの条件のもと選定された試験パラメータとアルゴリズムにより選定された結果における再現率を算出した。

4. 検証結果

4.1.1 論理構造同定

論理構造同定において可視化した結果を表1に示す。

表1 可視化結果

仕様	文数(文)	配置検数(個)	正答数(個)
pot-210	4	4	4
pot-220	6	6	5
pot-221	4	4	4
pot-230	3	3	3
pot-271	2	2	2
pot-312	2	2	2
pot-330	10	10	7
pot-500	10	10	8
合計		41	35

表の結果について pot-220 を例に説明する。実際に可視化した結果は図10である。図8よりパラグラフへ記述を変換した際には計6文で構成されており、要求文を含めた計6文が構造化の対象となる。正答数に関しては、3.2節で説明した人手による配置の結果と比較し、それぞれの文が同

じ構造上に配置されている場合には正しく配置されたと判断し、正答数としてカウントした。6文目は、人手での配置は「前件に強く補足説明」への配置だが、アルゴリズムによる結果は「含意関係の補足説明」へ配置されているため誤りとし、それ以外が人手と同じであるため、正答数は5である。表1よりアルゴリズムでの論理構造の同定は人手との結果と比較しその再現率は約85%となった。

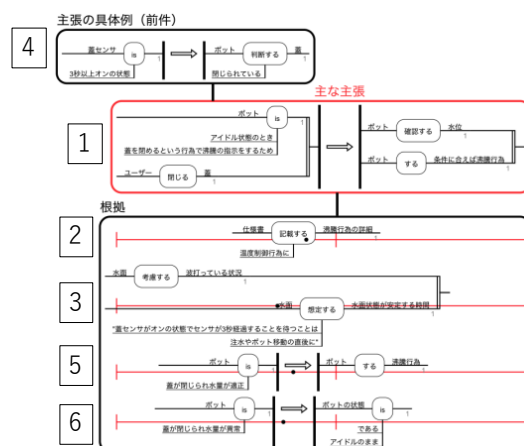


図10 アルゴリズムによる構造の可視化例

論理構造の同定に関して、再現ができなかった部分の例として pot-330 を例に説明する。図11は入力文と構造化した可視化結果を抜粋したものである。セミ形式記述に変換した際に前件となる部分が赤字であり、後件となるのが青文字である。8文目は後件の「温度制御行為はしない」という部分に対する追加要素として、「保温ランプ・沸騰ランプが共に消灯する」という内容を記載している。そのため、構造化した際には「後件に強く補足説明」されている要素への配置が好ましく、今回のような「含意関係の補足説明」へ配置されるものは誤りとしている。

- [1] 温度制御行為ができないポットの状態ではヒータをon/offさせると危険なため、ポットが沸騰行為ができず、ポットが保温行為ができないならば、温度制御はしない。
[8] 温度制御停止中の表示は、保温ランプ・沸騰ランプが共に消灯する。

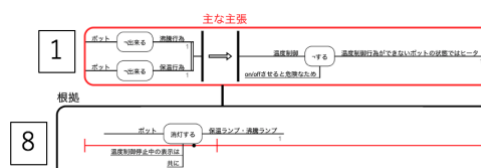


図11 pot-330の構造化

4.1.2 パラメータ抽出

試験パラメータ抽出支援に関しては、3.2節で説明したルールをもとに作成した試験パラメータと関連性の高い命題プリミティブの抽出結果を表2に示す。

表 2 試験パラメータ抽出結果

仕様	パラメータ選定数	正答数 (個)
pot-210	2	2
pot-220	3	2
pot-221	3	2
pot-230	2	2
pot-271	2	1
pot-312	1	1
pot-330	7	2
pot-500	5	1
合計	25	13

pot-220 より表の結果について説明していく。図 12 は試験パラメータとして内容が記載されている命題プリミティブを人手で色付けした結果である。

3.2 節より前件に関しては、青色で囲まれた「蓋を閉める」の部分に対し「蓋センサーが 3sec 経過 (2 文目, 4 文目)」, 後件のオレンジ色で囲まれた「条件下」の部分に対し「水量が適正 (5 文目)」, 緑色で囲まれた「沸騰行為」の部分に対し「3 章に詳細が記載されている旨 (2 文目)」の 3 つの内容を試験パラメータとして抽出を行った。

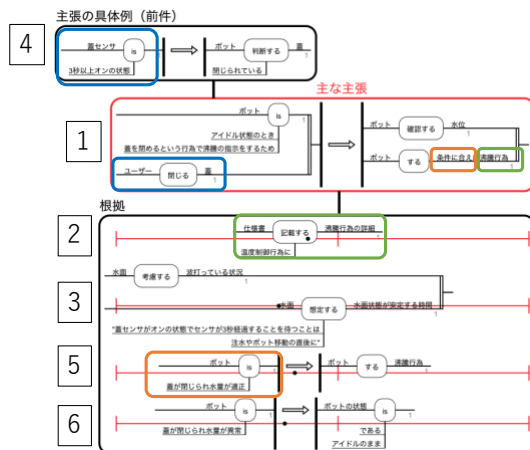


図 12 試験パラメータとなる情報の配置箇所

2.2.5 節で説明した組み合わせの B に着目した出力結果が表 3 である。青文字の部分試験パラメータとして考慮すべき内容であり、最上位に現れることで関連性が高いことが視覚的に把握できる。このような状態の時試験パラメータの支援が可能であるとし正答数をカウントする。

表 3 プリミティブの類似性

文番号	類似度	関係語	主体	対象	制約
1		「[w]」	「[ポット]」	「[.]」	「[アイドル状態のとき]」「蓋を閉める」という行為で沸騰の指示をするため
4	5.225	「[w]」	「[蓋センサー]」	「[.]」	「[3秒以上オンの状態]」
4	3.666666667	「[判断する]」	「[ポット]」	「[量]」	「[量]にわたっている」

次に組み合わせ D に着目した出力結果が表 4 である。

表 4 プリミティブの類似性

文番号	類似度	関係語	主体	対象	制約
1		「[する]」	「[ポット]」	「[条件に合えば沸騰行為]」	
2	9.261363636	「[記憶する]」	「[仕様書]」	「[沸騰行為の記録]」	「[温度制御行為に]」
3	8.681660823	「[決定する]」	「[水面]」	「[水面状態が安定する時間]」	「[「蓋センサーがオンの状態」でセンサーが3秒経過することを待つことは、注水やポット移動の直後に]」
5	4	「[する]」	「[ポット]」	「[沸騰行為]」	
6	3.512698413	「[w]」	「[ポットの状態]」	「[.]」	「[である]」「[アイドルのまま]」
3	3.169230769	「[考慮する]」	「[水面]」	「[波打っている状況]」	
5	1.688311688	「[w]」	「[ポット]」	「[.]」	「[蓋が閉じられ水量が適正]」
6	1.20952381	「[w]」	「[ポット]」	「[.]」	「[蓋が閉じられ水量が異常]」

緑色の文字は試験する際の考慮事項として抽出したい内容

であり、最上位で出力されている。しかし、オレンジ色に関しては、先程の青色文字や緑色文字とは異なり、最上位ではなく下位で出力されている。これらの状態の時試験パラメータが支援可能であるとは判断せず正答数としてはカウントしない。

このような判定のもと実施した結果において表 2 より、人手での抽出とアルゴリズムでの抽出における再現率としては、52%であった。

5. 考察

パラグラフ構成文の論理構造の同定に関しては、人手での配置に対し、アルゴリズムによるその再現率が約 85%であった。しかしながら、図 11 の再現ができなかった例を見ると、後件の「温度制御停止」に関して、8 文目で温度制御停止中の追加イベントとして「保温ランプ・沸騰ランプ共に消灯する」と後件にのみかかるような内容であるが、現状では「含意関係の補足説明」として配置されてしまう。これら誤配置の要因として、以下が考えられる。まず、しきい値の問題である。これは可視化モデルへ配置する際には図 6 に示した配置ルールをもとに行われているが、このパラメータの調整が現状では十分でない可能性がある。本アルゴリズムを多くの仕様書文に適用することにより、パラメータの調整を図っていく必要がある。

また、形態素解析により「温度制御停止」という複合語は「温度・制御・行為」と 3 つに分割され複合語としては扱っていないことが課題と考えられる。これにより、共起語候補が、複合語を構成することで、限定された範囲よりも大きく定義されてしまうことが要因ではないかと考える。

試験パラメータとなりうる要素の発見に関しての再現率は 5 割程度とあまり高い結果を得ることができなかった。

上記要因としては、命題プリミティブごとに類似性を算出する際にその文量を考慮していないことが挙げられる。例えば表 4 において、文番号 3 の類似度は約 8.7 あり、関連性のある文番号 2 に次いで高くなっている。今回試験パラメータ抽出において「類似度が高くなる＝文の関連性が高くその情報量も多い」と考えており、それぞれの単語の類似度を加算した結果を用いている。その結果文量が多い内容が上位を占めるかたちとなってしまっている。そのため、加算する類似度に関しては一定の閾値を設け、類似度の低い単語が数多く存在することで、合算した際の類似度が分量により大きくなる課題に対処する必要がある。

6. まとめ

本研究では、仕様書パラグラフ構成文の論理構造を同定し、試験パラメータの導出支援を目的としてアルゴリズムの検討とその検証を実施した。構造化やパラメータ抽出支援で

活用する文間の類似性を算出するために、単語の意味的類似性とコンテキストの類似性に着目した。その結果、論理構造の同定に関しては、人手との配置に対する再現率が約85%と高い結果となった。一方で、この構造化のもと実施したパラメータの導出支援に関しては、人手抽出に対する再現率が52%とあまり高い結果を得ることができなかった。この要因としては形態素解析した結果、複合語が分割され限定される範囲より多くの語で類似性が算出されることや、プリミティブの文量に差があることで合算される単語数が異なり、文量の多いプリミティブほど類似性が高く

算出されることに起因すると考える。

今後の予定としては、論理構造の同定の精度をより高めるため、本アルゴリズムをより多くの仕様文に適用しパラメータを調整していく必要がある。一方でパラメータ導出支援に関しては、複合語の扱いやプリミティブの文量等いくつかの要因が重なっている可能性が存在するため、要因の特定を行い、精度を向上することができるよう努めていきたい。

参考文献

- [1] 青山裕介, 黒岩丈瑠, 久代紀之. “テストケース生成のためのシステム仕様書の論理記述変換アルゴリズム.” 情報処理学会論文誌 61.3 (2020): 521-534.
- [2] Hitoshi Isahara, Francis Bond, Kiyotaka Uchimoto, Masao Utiyama and Kyoko Kanzaki Development of Japanese WordNet. In LREC-2008, Marrakech.
- [3] 黒橋禎夫, 自然言語処理, 放送大学教材, 2020
- [4] 京都大学黒橋・河原研究室, “日本語形態素解析システム JUMAN++”, <http://nlp.ist.i.kyoto-u.ac.jp/?JUMAN%2B%2B>
- [5] 緒方佑亮, 青山裕介, and 久代紀之. “仕様書におけるパラグラフ内構成文の論理構造の可視化.” 研究報告ソフトウェア工学 (SE) 2021.17 (2021): 1-7.
- [6] Kirschner, P.A., Buckingham-Shum, S.J., and Carr, C.S. (2012): “Visualizing argumentation: Software tools for collaborative and educational sense-making”, Springer Science & Business Media.
- [7] 話題沸騰ボット (GOMA-1015 型) 要求仕様書 [第 7 版], SESSAME, 2005.