

特許調査を効率化する請求項理解支援機能

小西 一也 新海 正吾 高木 徹 三部 靖夫

(株)NTT データ 技術開発本部
〒104-0033 東京都中央区新川 1-21-2

近年、多くの企業が特許活用の重要性に対する認識を高めている。特許活用において特許調査は必須の業務であるが、煩雑で時間がかかるという問題がある。例えば、特許請求の範囲が定義されている請求項は、一般的に長大で文構造が複雑な一つの複文であるため、理解することは容易ではない。従来、特許調査者は、特許請求の範囲を理解するとき、請求項を何度も読み返して発明特定事項を抽出し、それぞれの意味属性と発明特定事項間の関係を解釈して、特許請求の範囲を理解するために必要な情報を紙面に書き出す作業を実施していた。このような煩雑な作業が、特許請求の範囲の理解にかかる時間を長大化する原因となっていた。この問題を解決するために、我々は請求項理解支援機能を開発した。本機能は、請求項テキストを入力とし、請求項を構成する要素をノード、各要素間の関係をエッジとしたグラフを自動作成して出力する。本機能により、特許請求の範囲の理解に必要な情報が自動的に抽出され、特許調査者が特許請求の範囲を理解するための作業が効率化される。

A Support Function of Understanding Patent Claims for Efficient Patent Investigation

Kazuya KONISHI Shogo SHINKAI Toru TAKAKI and Yasuo SAMBE

Research and Development Headquarters, NTT DATA CORPORATION
1-21-2 Shinkawa, Chuo-ku, Tokyo, 104-0033, Japan

Recently, the importance of patent utilization has gradually been recognized by many corporations. Patent investigation is a necessary task for the application of patent, and so on; however, it is a laborious and time-consuming process. For example, since patent claim that defines inventions is generally written in a very long complex sentence, it is not easy for patent investigator to understand the patent claim. Traditionally, for understanding the patent claim, patent investigators have repeatedly read it and extracted claimed elements of the invention from the patent claim. Then, they have interpreted semantic attributes of each claimed elements of the invention and relation among them, and jotted them down on the paper as necessary information to understand the patent claim. This laborious operation was leading to the time-consuming process of understanding the patent claim. To solve this problem, we have developed a support function of understanding patent claims. The input to the function is claim text for which the investigator must understand, and the function automatically outputs the graph in which the element constituting the patent claim is used as node and relation between each element is used as edge. As a result of this function, necessary information for understanding the patent claim is extracted automatically and simplifies the process for patent investigators to understand the patent claim.

1. はじめに

近年、企業において特許活用の重要性に対する認識が高まっている。企業が利益を生み出す資源となる発明を創造し、特許化するためには、様々な場面で特許調査が必要になる。例えば、適切な研究開発テーマの選定のために技術動向を調査する場面や、発明の権利化可能性を高めるために先行例を調査する場面などがある[1]。2006年1月に特許庁が発表した「特許審査迅速化・効率化のための行動計画」¹では、企業に依頼内容の事前チェック徹底が要請されており、先行例調査は特に重要性が高いと言える。

特許調査では、特許明細書に書かれた発明を理解する必要がある。しかし、特許明細書の中で特許請求の範囲が定義されている請求項は、一般的に長大で構造が複雑な複文であるために読みにくく、理解には長い時間がかかるという問題がある。そこで本稿では、特許調査を効率化するために、特許調査者による特許請求の範囲の理解を容易にする、請求項理解支援機能を提案する。

2章では特許調査業務を2つの作業に整理して各作業の問題点を説明する。3章では特許請求の範囲の理解に関する問題点を解決する、請求項理解支援機能を提案する。4章では請求項理解支援機能を実現する請求項解析技術と現状の解析精度を説明し、最後に5章でまとめる。

2. 特許調査の問題点

我々は、特許調査業務を2つの作業に整理した。どちらの作業も煩雑であり、時間がかかるという問題がある。各作業の概要と主な問題点を以下に示す。

(1) 特許明細書の検索：

調査したい特許明細書を検索する。IPCなどの技術分野の指定による検索対象の絞り込みや、調査したい特許の内容を特徴的に表す検索語の指定による検索を実施する。技術分野を選定するときには、膨大な技術分野分類表を調べる必要がある。また、検索語を選定するときには、検索漏れや誤検索の抑止を十分に考慮する必要がある。し

¹ http://www.jpo.go.jp/torikumi/puresu/pdf/jinsokuka_patent/02.pdf

請求項テキスト

…、前記開閉部を開閉駆動する電動モータと、前記開閉部が全閉状態となったことを検出する全閉検出手段と、…

発明特定事項1

請求項構成要素1-1 (作用)	請求項構成要素1-2 (発明を特徴付ける 構成要素)
前記開閉部を開閉駆動する	電動モータと、

発明特定事項2

請求項構成要素2-1 (作用)	請求項構成要素2-2 (発明を特徴付ける 構成要素)
前記開閉部が全閉状態となったことを検出する	全閉検出手段と、

図 1. 発明特定事項と請求項構成要素の例

たがって、技術分野の選定、および検索語の選定を実施する場合、技術分野の全体的な体系を把握しておく必要があり、選定すべき情報の調査には時間がかかる。

(2) 特許請求の範囲の理解：

特許請求の範囲を読解する場合、一般的には各請求項を発明特定事項(特許法36条5項の「発明を特定するために必要と認める事項」)ごとに分割し、各発明特定事項間の関係を明確化することにより、発明の構成を理解する。請求項は長大で構造が複雑な1つの複文で記述されるため、発明特定事項の特定は容易ではない。このため、特許請求の範囲の理解には時間がかかる。

(1)特許明細書の検索に対しては、NTCIR²などにおいて多くの研究開発がされている。我々もこれまでに、調査対象となる特許明細書から類似特許検索に適した検索語を抽出し、類似特許検索を高精度化する方法を提案している[2][3]。

(2)特許請求の範囲の理解について、従来、特許調査者は紙面に印刷した請求項を何度も読み返しながら複数の発明特定事項を漏れなく抽出し、各発明特定事項について、発明の前提事項や特徴的事項といった位置付けを解釈して、紙面に書き込むなどの作業を実施していた。そして、特許調査者は自ら書き込んだ情報を参照し、特許請求の範囲を理解した。発明特定事項の抽出や位置付けの解釈は、通読の反復と熟考を要する非常に煩雑な作業であり、特許請求の範囲の理解にかかる時間を長大化する原因となっていた。

² <http://research.nii.ac.jp/ntcir/index-ja.html>

表 1. 請求項構成要素の意味属性

請求項構成要素の意味属性	付与する請求項構成要素の表現
発明の名称	発明の名称.
発明の動作特徴	発明を特徴付ける動作.
発明を特徴付ける構成要素	発明を特徴付ける構成要素. 通常は要素の名称.
発明の前提を特徴付ける構成要素	発明の前提を特徴付ける構成要素. 通常は要素の名称.
発明の前提の動作特徴	発明を前提を特徴付ける動作.
参照請求項	参照請求項.
参照請求項の内容	参照請求項の内容. 通常は参照請求項における発明の名称.
作用	発明を特徴付ける構成要素が有する作用.
前提表現	発明特定事項が前提的なものであることを示す表現。「において」「であって」など.
構成表現	発明の構成を示す表現。「で構成された」「を備える」など.
特徴表現	発明の特徴を示す表現。「ことを特徴とする」など.
その他	その他, 発明特定事項の作用を限定する条件など.

3. 請求項理解支援機能

3.1. 請求項構成要素

我々は、特許調査者による特許請求の範囲の理解にかかる時間を短縮することを目的として、発明特定事項の抽出や位置付けの解釈という煩雑な作業を、テキスト処理により自動化する、請求項理解支援機能を開発した。

ここで、図1の例を用いて、本機能が扱う請求項構成要素を説明する。発明特定事項には、請求項特有の様々な意味属性を有した情報がある。この情報を抽出したものを、本稿では請求項構成要素と呼ぶ。図1に示す2つの発明特定事項は、それぞれ「発明を特徴付ける構成要素」という意味属性の情報と、「作用」という意味属性の情報を含む。我々は、請求項特有のこれらの情報を区別して特許調査者に示すことで、発明特定事項の内容の理解を効率化できると考え、これらの情報を請求項構成要素として抽出することを考えた。

請求項構成要素とその意味属性の定義として、請求項記述言語 PCML が考えられる。PCML は、知的財産マネジメント研究会の特許戦略工学分科会³において、請求項のコンピュータ処理のインフラとして検討されている XML フォーマットである。PCML の定義を、請求項構成要素とその意味属性の定義とすることは、本機能の普及に有効である。しかし、定義は検討中の

ものであり、実際に請求項に書かれている情報と各要素との対応が完全ではない。そこで、我々は独自に請求項構成要素を定義した。請求項構成要素の意味属性の一覧を表1に示す。この定義は、多くの請求項を精読して人手で請求項構成要素を抽出し、各請求項構成要素に明確に付与できる意味属性を検討した結果である。

3.2. 機能説明

請求項理解支援機能は、請求項テキストを入力とし、請求項構成要素をノード、請求項構成要素間の関係をエッジとしたグラフを自動作成して出力する。特許調査者は、発明特定事項の抽出や位置付けの解釈という煩雑な作業を実施することなく、グラフを目視することで、構成を認識しながら請求項を読解できる。また、請求項理解支援機能の別の出力形式として、請求項構成要素の意味属性ごとのリスト表示や、請求項をノード、請求項間の参照関係をエッジとしたグラフの表示もある。これらの出力も、特許請求の範囲の理解を容易にするために重要な情報である。さらに、解析結果表示画面において、特許調査者によるグラフの編集と、編集結果の共有を可能にしている。この機能は、複数人の特許調査者による特許請求の範囲の理解を効率化するものである。

3.2.1. 請求項解析機能

請求項解析機能は、請求項構成要素を抽出してそれぞれに意味属性を付与し、さらに請求項

³ http://www.patentisland.com/Patent_Strategy_Engineering/Result.html

特開2006-009474号【請求項1】	
意味属性	請求項構成要素
発明の前提の動作特徴	車両に設けられた開閉部材を自動的に開閉する
発明の名称	車両用自動開閉装置
前提表現	であって、
作用	前記開閉部材を開閉駆動する
発明を特徴付ける構成要素	電動モータと、
作用	前記開閉部材が全閉状態となったことを検出する
発明を特徴付ける構成要素	全閉検出手段と、
作用	前記開閉部材の移動に応じてパルス信号を出力する
発明を特徴付ける構成要素	パルス信号出力手段と、
作用	前記開閉部材が基準位置にあるときを起点として前記パルス信号を推算することにより前記開閉部材の位置を検出する
発明を特徴付ける構成要素	位置検出手段と、
作用	前記位置検出手段により検出される前記開閉部材の位置に基づいて前記電動モータの作動を制御するとともに前記全閉検出手段により前記開閉部材が全閉状態となったことが検出され且つ前記開閉部材が全閉位置にあることが検出されたときに前記開閉部材の全閉状態を認識する
発明を特徴付ける構成要素	制御手段
構成表現	とを有し、
作用	前記パルス信号出力手段が前記パルス信号を出力することができない状態となったときには
発明の動作特徴	前記位置検出手段に前記開閉部材が全閉位置にあると認識させる
特徴表現	ことを特徴とする
発明の名称	車両用自動開閉装置。

特開2006-009474号【請求項2】	
意味属性	請求項構成要素
参照請求項	請求項1記載の
参照請求項の内容	車両用自動開閉装置
前提表現	において、
発明を特徴付ける構成要素	前記開閉部材に設けられたラッチが前期車両に設けられたストライカに完全に係合したことをフルラッチスイッチが検出し、
作用	前記ラッチにラチェットに係合したことをラチェットスイッチが検出したときに
発明を動作特徴	前記制御手段が前記開閉部材の全閉状態を検出する
特徴表現	ことを特徴とする
発明の名称	車両用自動開閉装置。

図 2. 請求項構成要素の例

(特開 2006-009474 号 請求項 1 および請求項 2)

構成要素間関係の関付けを行う機能である。

図 2 に、特開 2006-009474 号の公開特許公報の請求項 1 および請求項 2 から、請求項構成要素を抽出して意味属性を付与した結果を例示する。さらに、意味的に関連のある請求項構成要素同士を関係付け、参照関係にある請求項同士を関係付ける。これらの方法により、次節で説明する解析結果表示機能を実現する。

なお、請求項解析を実現する技術は 4 章で説明する。

3.2.2. 解析結果表示機能

解析結果表示機能は、請求項解析機能の出力結果を、以下の 3 形式で表示する機能である。

(1) 請求項構造表示形式：

請求項解析結果を、請求項構成要素をノードとし、請求項構成要素間の関係をエッジとしたツリーグラフで表示する。図 3 に表示例を示す。この表示では、上下の請求項構成要素の並びが、請求項の記述順に合致しているため、並び通りに請求項構成要素を読み進めることで、請求項本文を読解することができる。グラフ上では、読んでいる部分の意味属性を意識しながら請求項を読めるため、請求項の理解が効率的になる。例えば図 3 の表示例で請求項構成要素を読み進め、上から 3 つ目の請求項構成要素「前記開閉部材を開閉駆動する」という文字列を読むときには、それが発明を特徴付ける構成要素の 1 つである「電動モータ」の作用であることを意識しながら読むことができる。

(2) 請求項構成要素リスト表示形式：

請求項から抽出した請求項構成要素を、意味属性ごとに関数してリスト表示する。図 4 に表示例を示す。表示例を一見するだけで、この発明には前提となる記述があり、5 つの発明を特徴付ける構成要素を有し、発明の動作特徴が示されていることがわかる。この表示を確認することで、発明の動作特徴のみを抽出するなど、特定の意味属性の請求項構成要素の発見が容易になる。

(3) 請求項間関係表示形式：

請求項間の参照関係をツリーグラフで表示する。図 5 に表示例を示す。表示において、参照関係は矢印で表す。点線の矢印は、参照元の請求項に含まれる発明を特徴付ける構成要素の作用を限定することを示し、実線の矢印は、参照元の請求項に新たに発明を特徴付ける構成要素を追加することを示す。表示例から、独

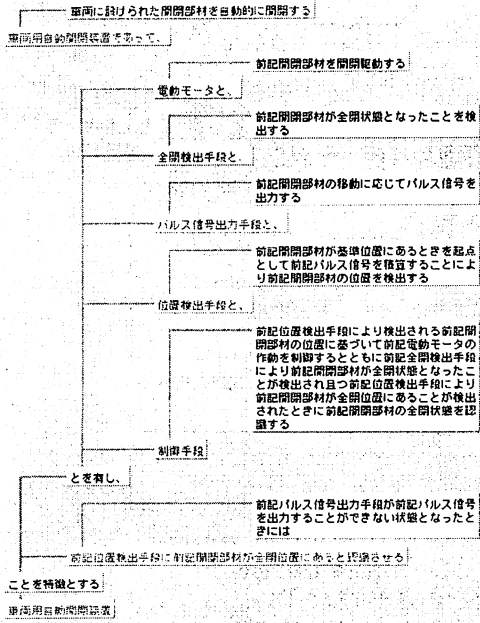


図 3. 請求項構造表示の例

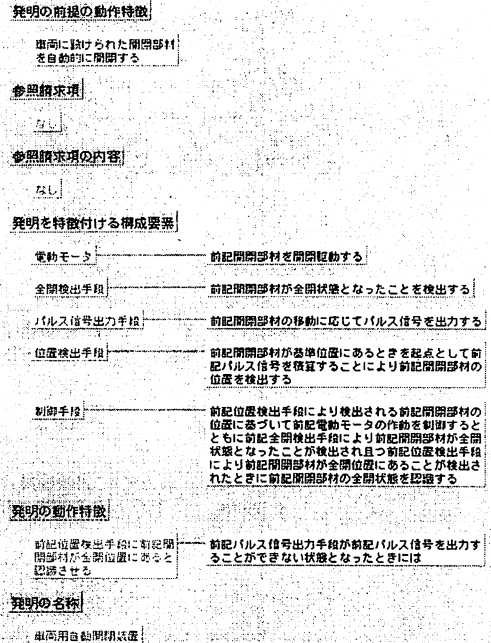


図 4. 請求項構成要素リスト表示の例

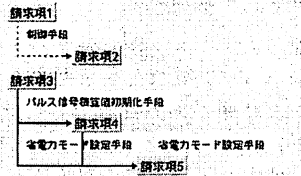


図 5. 請求項間関係表示の例

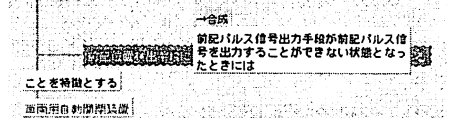


図 6. 解析結果編集(請求項構成要素の合成)の例

立請求項として請求項 1 と請求項 3 があり、請求項 2 は請求項 1 の従属項、請求項 4 および請求項 5 は請求項 3 の従属請求項であることがわかる。また、請求項 2 は請求項 1 の制御手段を限定し、請求項 4 は請求項 3 にパルス信号積算値初期化手段を追加した構成であることがわかる。この表示を閲覧することで、独立請求項の抽出や、従属請求項を読解する前の予備知識を得ることができ、請求項の理解が効率的になる。

3.2.3. 解析結果編集・共有機能

解析結果編集・共有機能は、特許調査者による請求項解析結果の任意の編集と、複数人の特許調査者での編集結果の共有を実現する機能で

ある。具体的には、解析結果表示画面において、請求項構成要素の合成や分割、意味種別の変更、メモ付与といった編集と、編集結果の保存を容易に実施でき、保存された編集結果を複数人の特許調査者が閲覧できる。複数人の特許調査者で1つの特許を調査する場合、この機能により特許調査者同士で理解内容を確認し合うことで、特許請求の範囲の理解をさらに効率化できる。

発明特定事項が階層的に構成されているような複雑な発明の場合、特許請求の範囲の記述があいまいになり、特許調査者ごとに発明特定事項の解釈が異なる場合が考えられる。例えば、「スイッチを有する制御装置を具備しさらに熱で発電する発電装置を備えた非常用電源」という記述の請求項があった場合、制御装置を具備するものが、発電装置か非常用電源か、明確ではなくなる。そこで、複数人の特許調査者が独自の解釈で解析結果を編集したり、解釈の観点をメモ付与したりして保存し、互いに確認し合

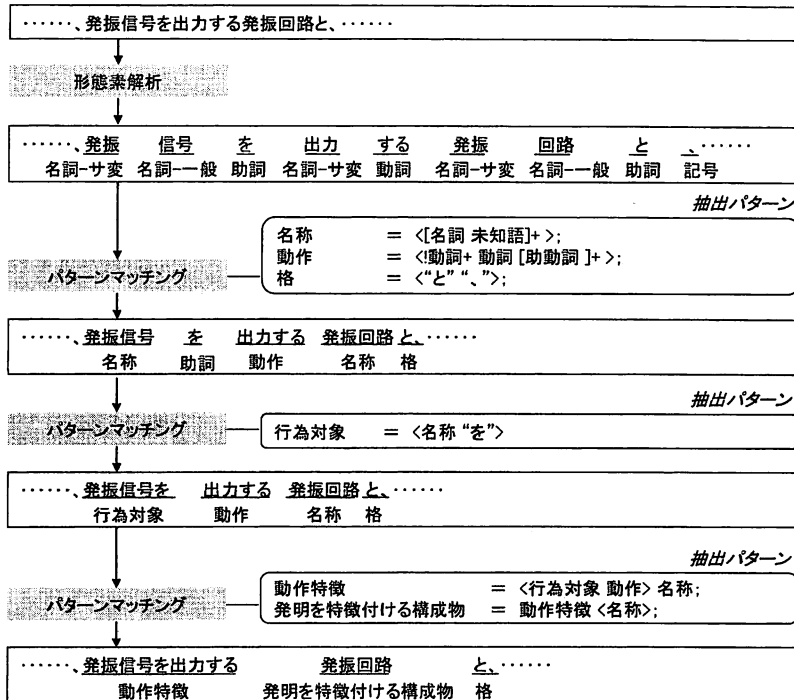


図 7. パターンマッチングによる請求項構成要素抽出の例

うことで、特許請求の範囲の理解をさらに効率化できる。

図 6 に、解析結果編集のイメージを示す。これは請求項構成要素を合成するイメージであり、請求項構成要素を合成したい別の請求項構成要素の上に、マウス操作でドラッグした状態を示している。このままドロップするだけで、2つの請求項構成要素を合成できる。他の解析結果編集も、容易な操作で実施できるようにしている。

4. 請求項解析技術

4.1. パターンマッチングによる請求項解析

これまでに、請求項解析を実現する技術として、手がかり句を用いた特許請求項の修辞構造解析技術が研究されている[4]。この技術によっても、請求項構成要素の抽出と請求項構成要素間の関係付けを実施でき、結果をグラフ表示することができる。しかし、各請求項構成要素に意味属性を付与しないため、請求項構成要素リスト表示や請求項間関係表示への適用は難しい。

この他に、テキストから特定パターンの形態素列を抽出して意味属性を付与する、パターンマッチング技術が開発されてきた[5]。これは、テキストと抽出パターンを入力として、テキスト中から抽出パターンに一致する文字列を抽出し、意味属性を付与して出力する、情報抽出技術である。抽出パターンは、左辺に意味属性、右辺に形態素や形態素列で構成したパターンを指定する。本稿では、パターンを構成する形態素や形態素列を示す要素を、パターン要素と呼ぶ。パターンマッチングでは、抽出パターンの右辺に指定されたパターンに合致する文字列をテキスト中から探索し、合致した文字列を抽出して、左辺に指定された意味属性を付与する。文献[5]のパターンマッチング技術では、パターン要素に形態素の表層語だけでなく、品詞なども指定できるようにしているため、作成するパターンを少なくできる。また、パターンマッチングで付与した意味属性を、さらにパターン要素に指定できるようにしているため、複雑なパターンの作成が容易になる。

請求項解析技術では、請求項特有の言い回しを反映した抽出パターンを作成し、文献[5]のバ

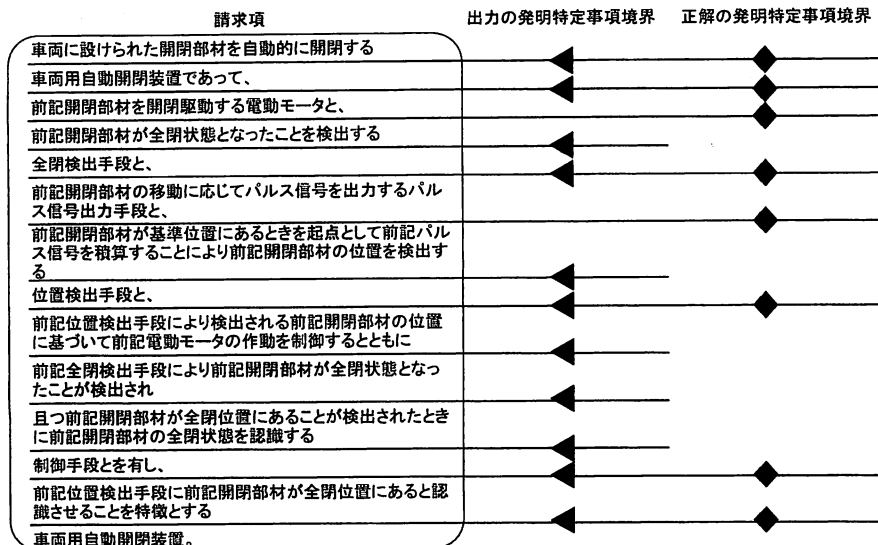


図 8. 発明特定事項の境界の例

ターンマッチング技術を用いて、前章に述べた請求項構成要素を抽出する。請求項構成要素を抽出する具体的な処理内容を、図 7 を用いて説明する。まず、入力となる請求項テキストを形態素解析し、形態素に分割して各形態素の品詞を特定する。次にパターンマッチングを実施し、抽出パターンに合致する形態素列に意味属性を付与する。例えば、名詞や、形態素の辞書に登録されていない未知語が連続する形態素列に「名称」という意味属性を付与する。次いで、意味属性が「名称」である形態素列に、表層語が「を」である形態素が連続する形態素列を抽出して、「行為対象」という意味属性を付与する。このように、指定されたすべての抽出パターンを用いたパターンマッチングを順次実施し、最終的には請求項構成要素に相当する形態素列を抽出する。

パターンマッチングで請求項構成要素を抽出して意味属性を付与した後、意味属性の並びのパターンに従い、請求項構成要素間の関係付けを実施する。例えば、意味属性が「発明を特徴付ける構成要素」である請求項構成要素 A の直前に、意味属性が「作用」である請求項構成要素 B がある場合、B は A の作用を示すものと判定し、A に関係付ける。これにより請求項構成要素の抽出と意味属性の付与、および請求項構成要素間の関係付けを実現でき、請求項構造表示形式、請求項構成要素リスト表示形式の解

析結果表示が可能になる。

また、各請求項について、意味属性が参照請求項である請求項構成要素を用いて、参照関係にある請求項間を関係付ける。これにより、請求項間関係表示形式の解析結果の表示が可能になる。

4.2. 請求項解析精度の評価

我々は、請求項構成要素の特定および意味属性の付与のために 241 の請求項構成要素抽出パターンを作成し、請求項解析精度を評価した。評価は、発明特定事項分割の正しさについて実施した。これは、請求項構成要素は特許調査者による特許請求の範囲の理解をより容易にするための詳細な情報であり、特許調査者が発明の構成を理解するために最終的に特定すべきものは発明特定事項であることによる。

まず、2002 年公開特許公報のうち、筆頭 IPC セクション A~H(C を除く)それぞれから 100 件ずつ、合計 700 件をランダムに抽出し、各公報の第 1 請求項について人手で発明特定事項の分割を実施した。この分割結果における発明特定事項間の境界を「正解の発明特定事項間境界」とする。次に、同じ 700 件の請求項を対象として請求項解析を実施し、解析結果における請求項構成要素を、一定の規則に従って結合した。この規則は、請求項構成要素間の関係付け規則

の一部と同一であり、関係付けられる請求項構成要素同士を結合するものである。この処理により結合した請求項構成要素を発明特定事項とみなし、境界を「出力の発明特定事項間の境界」とする。図8に、請求項を発明特定事項に分割したときの境界の例を示す。図中の三角印は出力の発明特定事項境界、菱形印は正解の発明特定事項境界を示している。最後に、正解と解析結果の発明特定事項間境界を比較し、適合率、再現率、F値を算出した。各値の意味と算出式を以下に示す。

・適合率(P)：

出力の発明特定事項間境界について、誤りの少なさを表す指標

$$P = \frac{\text{正解と一致する出力境界の数}}{\text{すべての出力境界の数}}$$

・再現率(R)：

出力の発明特定事項間境界について、漏れの少なさを表す指標

$$R = \frac{\text{正解と一致する出力境界の数}}{\text{すべての正解境界の数}}$$

・F値(F)：

適合率と再現率の調和平均

$$F = \frac{2 * P * R}{P + R}$$

図8の例では、出力の境界が11箇所、正解の境界が8箇所あり、6箇所が一致しているため、適合率が55%(6/11)、再現率が75%(6/8)となる。

700件の請求項に対して請求項解析を実施した結果は、適合率78%、再現率87%、F値82%であった。8割程度の解析精度は、特許調査者による任意の解析結果編集が可能であることも加味し、請求項理解支援機能への適用が可能であることを示していると考えられる。また、解析ルールを追加することで、一部の解析誤りを修正でき、解析精度を向上できる見込みも確認している。

5. まとめ

本稿では、請求項から請求項構成要素を抽出して意味属性を付与し、請求項構成要素間の関係付けを実施して結果をグラフ表示する、請求項理解支援機能を提案した。本機能は、特許調査における特許請求の範囲の理解において、従来、特許調査者が発明特定事項を特定するため

に実施していた煩雑な作業を自動化するものである。

参考文献

- [1] 谷川, 河本, “特許工学入門”, pp.61-80, 中央経済社, 2003.
- [2] 小西, 北内, 高木, “発明の特徴に着目した検索語抽出による先願特許検索”, 第15回データ工学ワークショップ(DEWS2004)論文集, 3-B-01, <http://www.ieice.org/iss/de/DEWS/proc/2004/paper/3-B/3-B-01.pdf>, 2004.
- [3] Konishi, Kitauchi, Takaki, "Invalidity Patent Search System of NTT DATA," Proceedings of the Fourth NTCIR Workshop on Research in Information Access Technologies Information Retrieval, Question Answering and Summarization, <http://research.nii.ac.jp/ntcir/workshop/OnlineProceedings4/PATENT/NTCIR4-PATENT-KonishiK.pdf>, 2004.
- [4] 新森, 奥村, 丸川, 岩山, “手がかり句を用いた特許請求項の構造解析,” 情報処理学会論文誌, Vol.45, No.3, pp.891-905, 2004.
- [5] 江里口, 木谷, “パターンマッチング手法による名称特定処理の有効性の検討,” 情報処理学会研究報告, NL-115-10, pp.9-15, 1996.