

# 特許明細書の作成効率・品質を大幅に向上する特許明細書半自動生成システム

Patent specification semiautomatic generation system for semi-automatic generation of specification from the claims.

IRD 国際特許事務所 所長・弁理士

谷川 英和

1986年神戸大学工学部システム工学科卒業。同年、松下電器産業(株)〔現パナソニック〕に入社し、中央研究所等において、データベース管理システム等の研究開発に従事。1999年弁理士試験合格。2002年1月、IRD国際特許事務所を開設。所長、弁理士。2003～2007年3月京都大学COE研究員、2007年4月～京都大学非常勤講師(現客員教授)、2011年4月～大阪大学非常勤講師(現招聘教授)。博士(情報学)。弁理士会、日本知財学会、情報処理学会各会員。2007年度から特許産業日本語委員会委員。

## 1 はじめに

従来、弁理士や企業の知財部門、研究開発部門においては、個人のスキルに頼った特許開発が行われていたため、高い生産性で特許明細書が作成されているとはいえなかった。また、特許明細書の品質の維持は、もっぱら個人のスキルに頼っており、品質管理の概念が浸透していないのが現状である。

そこで、本論文では、特許開発を効率化、高品質化するために、3種類の特許部品データベースを再帰的に検索し、特許明細書を構成する文章を取得し、特許庁フォーマットに従って、取得した文章を合成することにより、特許明細書を生成する特許明細書半自動生成システム(Patent Generator「以下、適宜、「PG」ともいう。))を提案する。

本稿の構成は以下のとおりである。2章では、関連研究・技術について述べる。3章では、特許部品データベースに基づく特許明細書生成エンジン(PG)について述べる。4章では、PGを用いて、コンピュータ・ソフトウェア関連発明の特許明細書を作成した際の評価結果について述べる。最後に5章では結論について述べる。

## 2 関連研究・技術

特許明細書の作成を支援するツールとして、特許ストーリー<sup>1)</sup>が存在する。

特許ストーリーは、過去の案文や公報等を保存しているフォルダを人手により指定して、文書を取込み、その

文書から必要な文をコピーすることにより、明細書の作成を効率化しようとするものである。

この既存ツールは、蓄えたノウハウや技術の実現手段についての文章をデータベース化して、利用するアプローチを採っていないので、特許明細書の品質向上には十分に寄与せず、作成効率の向上も不十分である。

## 3 特許明細書半自動生成システム (Patent Generator)

### 3.1 Patent Generator (PG) の概要

PGは、特許請求の範囲または明細書設計書(以下、明細書設計書等という。)を入力として、特許明細書を半自動生成するシステムである。

明細書設計書は、書類「特許請求の範囲」に加えて、特許明細書の設計のために3種類の情報を記載したドキュメントであり、特許明細書を記載するための指針となる成果物である。明細書設計書の例を、図1に示す。明細書設計書には、特許請求の範囲に加えて、以下の5種類の情報を記載する。

- ・技術開示実施形態番号

技術開示実施形態番号は、各請求項の技術開示を行う実施の形態の番号であり、請求項ごとに記載される。

- ・概要説明および特記事項

概要説明は、請求項の概要である。請求項の概要説明が実施の形態の概要となる。また、特記事項は、各請求項の技術開示を行うにあたって、特に触れたい事項である。特記事項は、請求項ごとに記載されるが、必須の項

<p>&lt;全体課題&gt;出願書類を効率的に作成できなかった          &lt;全体効果&gt;出願書類を効率的に作成できる          【請求項1】(1)(発明情報、説明情報に基づいて、出願文書情報を作成する)          1以上の発明に関する情報であり、発明を特定する情報である1以上の発明特定情報を有する情報である<b>発明情報</b>が記憶される</p> <p><b>発明情報記憶部</b>と、          発明情報に含まれる発明特定情報に関する説明を示す情報である説明情報が記憶される<b>説明情報記憶部</b>と、          前記発明情報記憶部が記憶している発明情報、及び前記説明情報記憶部が記憶している説明情報に基づいて、特許出願の文書に関する情報であり、発明情報に関する1以上の実施の形態を含む情報である<b>出願文書情報</b>を作成する<b>作成部</b>と、          前記作成部が作成した出願文書情報を入力する<b>出力部</b>と、を備えた出願文書情報作成装置。</p> <p>&lt;課題&gt;出願書類を効率的に作成できなかった          &lt;効果&gt;出願書類を効率的に作成できる</p> <p style="text-align: center;"><b>個別課題・効果</b></p>	<p><b>全体課題・効果</b></p>	<p><b>実施の形態番号 発明の概要</b></p> <p>クレーム、実施例設計書、クレームの構成要素を単に並べた情報など(発明ごとになっていなくてもよい)</p> <p><b>一時記憶可能</b></p> <p>特許出願できる状態であってもよく、その手前のたたき台であってもよい。          主に実施の形態の文書、効果や手段、課題などは、あってもなくてもよい</p> <p><b>用語の定義・説明</b></p>
--	-----------------------	--

図1 明細書設計書の例

目ではない。

・各請求項の文言詳細説明

請求項の文章中の用語の中で、特に、説明を要する用語が存在する。文言詳細説明は、その用語に対する説明である。説明を要する用語は、例えば、発明の説明のために用いた造語である。文言詳細説明は、図3の「コメント」である。

・発明全体の課題・効果

明細書設計書に、本願発明の全体の課題、および効果を記載しておけば、明細書および要約書に反映される。

・請求項ごとの課題・効果

明細書設計書に、請求項ごとの課題・効果を記載しておけば、明細書に反映される。

また、PGは、PGデータベース、およびPG処理部からなる。そして、PGデータベースは、特許庁フォーマット、と3種類の特許部品データベースからなる(図2参照)。

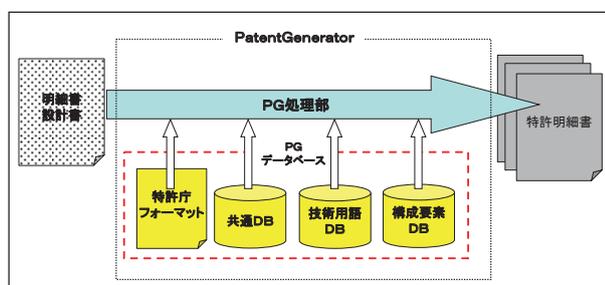


図2 PGの構造

特許庁フォーマットは、特許明細書、特許請求の範囲などの特許出願に必須の特許書類の記載項目を定義している。また、3種類の特許部品データベースとは、技術分野や製品の種類に依存しない文章等が格納された共通

データベース、発明の技術分野に依存する文章等が格納された技術用語データベース、構成要素の説明文を格納している構成要素データベースである。

さらに、PG処理部は、明細書設計書等を解析し、その解析の結果である発明の構成要素名、および特許庁フォーマットを構成するタグ名をキーとして、3種類の特許部品データベースを、順次、再帰的に検索し、文章を取得する。そして、取得した文章を、特許庁フォーマットに従って挿入し、特許明細書の一部を自動的に生成する。

広い範囲の技術開発を行い、毎年、多数の特許出願を行う大企業でPGを利用する場合、共通データベースと、技術用語データベースと、構成要素データベースの数の比は、1:n:m (1<=n<=m) である。

そして、このような企業において、多数の事業部門を統括する部署で共通データベースを構築し、各事業部門で技術用語データベースを構築し、各製品担当部門で構成要素データベースを構築する、という具合に、3種類のデータベースを分散開発し、かつ、特許明細書の作成者は、PGを用いて、統合的に3種類のデータベースを利用する。

したがって、PGデータベースが3種類の特許部品データベースを有することは有効である、と考える。

### 3.2 PGデータベース

PGデータベースを構成する3種類の特許部品データベースについて述べる。

#### (1) 共通データベース

共通データベースは、特許庁フォーマットのタグに対



応する定型文、または説明手順の情報を格納している。説明手順の情報とは、発明を説明する手順を定義する情報である。共通データベースの属性は、「タグ」および「タグ説明情報」である。「タグ」は、特許庁フォーマットで規定されているものであり、墨付き括弧（**Ⓜ**）で示される。「タグ説明情報」は、定型文や、説明手順を定義する情報が属性値となる。「タグ説明情報」は、予約語変数を含む。予約語変数は、PG で予め決められている変数であり、例えば、「発明の名称」や「構成要素」などである。予約語変数の値は、明細書設計書を解析して取得される。

## (2) 技術用語データベース

技術用語データベースは、技術分野（電気、機械、化学、コンピュータ・ソフトウェアなど）ごとに構築されるデータベースであり、技術用語の説明文や、その技術分野の発明の権利強化を図るための定型文や、文献としての役割を果たすための定型文や、技術分野固有の発明の説明手順の情報の集合である。技術用語データベースは、「説明項目」「定型文・説明手順」の属性を有する。「説明項目」は、主として技術分野固有の説明するべき項目名である。また、「定型文・説明手順」は、主として、「説明項目」に対応する定型文や、説明手順の情報である。定型文は、特許明細書の明瞭性や論理性を担保し、十分に技術開示を行い、発明が実施可能であるように記載しなければならない、という実施可能要件（特許法第 36 条第 4 項）を満たす意義や、特許の権利範囲を拡大する意義を有し、特許明細書作成におけるノウハウを表現した文章である。

## (3) 構成要素データベース

構成要素データベースは、発明の対象の製品ごとに構築されるデータベースであり、発明を構成する構成要素の実現手段の情報の集合である。例えば、コンピュータ・ソフトウェア関連発明の特許明細書において、記載しなければならない実現手段は、コンピュータおよび周辺機器を構成する部品の実現手段である。コンピュータおよび周辺機器の部品は、CPU や記録媒体やディスプレイなど、概ね決まっている。構成要素データベース中の情報は、権利範囲を拡大するための実現手段、または発明の成立性（特許法第 29 条第 1 項柱書）や実施可

能要件（特許法第 36 条第 4 項）といった特許要件を満たすための実現手段を定義したレコードであり、コンピュータ・ソフトウェア関連発明の特許を記載する際に、著者らが蓄えてきたレコードである。なお、部品の実現手段は、発明の成立性（特許法第 29 条第 1 項柱書）や実施可能要件（特許法第 36 条第 4 項）の特許要件を満たすために、通常、記載しなければならない。さらに、実現手段は製品により異なるので、例えば、発明の対象の製品がテレビであれば、テレビ用の構成要素データベースが存在し、エアコンであれば、エアコン用の構成要素データベースが存在することとなる。

## 3.3 Patent Generator の動作概要

PG を構成する PG 処理部は、明細書設計書等に記載された各請求項を自然言語解析し、請求項ごとに、発明の名称、1 以上の構成要素名、各構成要素の説明などの予約語変数の値、構成要素間の階層関係、および請求項数などの予約変数の値を取得する。

そして、PG 処理部は、請求項ごとに、明細書設計書等に記載された「技術開示実施形態番号」、「概要説明および特記事項」、「文言詳細説明」、「発明全体の課題・効」「請求項ごとの効果」の 5 種類の情報を取得する。そして、取得した「概要説明および特記事項」は、予約語変数<概要説明および特記事項>の値となる。なお、「技術開示実施形態番号」および「文言詳細説明」は、実施の形態における構成要素説明の文章生成に利用される。

次に、PG 処理部は、特許庁フォーマットを読み出す。そして、特許庁フォーマットに記載されたタグ（**Ⓜ**）をキーとして、共通データベースからレコードを検索し、検索したレコードのタグ説明情報を、タグの直後に追記する。そして、共通データベースのタグ説明情報が予約語変数を含む場合、既に得ている予約語変数の値をタグ説明情報に代入し、タグの直後に挿入する文章を生成する。

次に、共通データベースのタグ説明情報が一般変数を含む場合、技術用語データベースの検索処理を行い、一般変数の値を取得する。この処理を、特許庁フォーマットに含まれるタグ（**Ⓜ**）の数だけ繰り返す。

## 3.4 Patent Generator の他の機能

PG は、作成した明細書から符号の説明を自動生成す

る符号の機能を有する。また、PG は、作成した明細書から図面の簡単な説明を自動生成する機能を有する。

さらに、PG には、指定した範囲の文章から自立語を抽出し、当該自立語をキーとして、技術用語データベースおよび構成要素データベースを検索し、説明を補充する説明補充機能がある。

これらの機能も、面倒な作業を軽減し、明細書作成者の負担を低減するために有用である。

## 4 PG の評価

### 4.1 PG を利用した場合の作業効率向上

3 名の特許明細書作成者が、約 1 年間にわたり、PG のプロトタイプ版を使用し、約 100 件の特許明細書を作成した。PG を利用して作成した特許明細書の技術分野は、コンピュータ・ソフトウェア関連発明である。PG を利用した場合、特許明細書作成に要した作業時間は、約 15% 減少した。また、PG を用いて作成した特許明細書のうち、10 件の特許明細書を取り出して、PG による文章の生成率を算出した。その結果、平均 39.5% の情報が、PG により自動生成された。PG の文章の生成率が、そのまま作業時間の短縮に結びついていないのは、定型的文章や多数の特許明細書で共用される文章の作成が、発明特有の文章の作成ほど、労力がかからないからである、と考えられる。ただし、PG を用いて特許明細書を作成する場合、発明特有の文章の作成に注力すれば良いので、特許明細書作成の負担は、時間や量の効率化以上に減じられていると考える。

### 4.2 PG による特許明細書の品質向上

PG を用いて作成した特許明細書 10 件と、PG を用いずに作成した特許明細書 10 件の特許明細書品質を独自開発の特許明細書品質評価ツール<sup>2)</sup> を用いて算出した。PG を用いずに作成した特許明細書 10 件は、コンピュータ・ソフトウェア関連発明の分野の特許明細書であり、公開特許公報からランダムに抽出した。その評価結果を図 3 に示す。評価結果によれば、PG 利用の特許明細書品質が PG 非利用の特許明細書品質を大きく上回った。PG を利用して作成した特許明細書と PG を利用していない特許明細書とを比較して、特に、実施可能担保度と発明展開度において、高い数値が出ている。

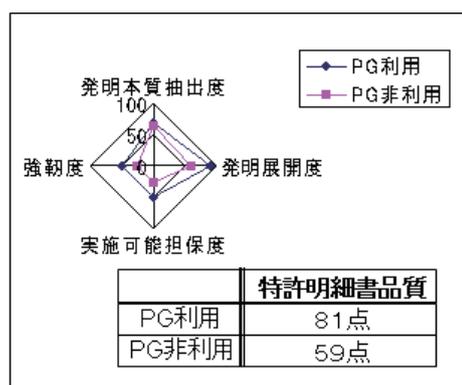


図 3 PG による特許明細書の品質向上

また、特許明細書の作成に熟練した弁理士などのノウハウを PG データベースに登録することにより、特許明細書作成の経験の浅い者でも、PG を利用して特許明細書を作成すればその品質は、一定以上になるものと期待できる。

## 5 おわりに

本論において、高品質な特許明細書を効率的に作成するための特許明細書生成エンジン (Patent Generator) について述べた。

さらに、Patent Generator により、特許明細書の作成効率が大幅に向上したこと、および特許明細書品質評価ツールを用いて、Patent Generator により特許明細書の品質向上が図られたと考えられることについても報告した。

### 参考文献

- 1) URL [http://www.ibr.co.jp/lessonpat/app\_a.html]
- 2) 谷川英和、新森昭宏、奥村学、難波英嗣「パテントメトリクスによる特許品質評価および特許価値評価」2006 年度 Japio 「特許情報活用の時代の検索と機械翻訳技術」、pp154-159、2006。