



## 2008 年度上期未踏 IT 人材発掘・育成事業(未踏ユース)採択案件評価書

### 1. 担当PM

竹内 郁雄 PM(東京大学大学院 情報理工学系研究科 創造情報学専攻 教授)

### 2. 採択者氏名

チーフクリエイター: 増満 光  
コクリエイター: なし

### 3. プロジェクト管理組織

株式会社 創夢

### 4. 委託金支払額

2,995,043 円

### 5. テーマ名

「特許請求の範囲」のモデル化と可視化及び入出力システムの開発

### 6. 関連Webサイト

なし

### 7. テーマ概要

近年、自社の発明に対して特許権を取得し、その権利を利用して他者との競争を勝ち抜こうとする動きがますます活発になってきた。そのため各者は、他者よりも早く自身の発明を権利化することが必要になってきている。特許を利活用するためには、その前提として、他の特許に対する調査を行う必要がある。特許調査とは、特許電子図

書館などで公開されている公開特許から、詳細を読み解くといった作業である。しかし、現在の特許調査は、特許権の利活用を行っている者と別の人物や機関が行っている場合が多い。これは、新規特許申請が年間 40 万件を超えることや発明の詳細を記した「特許請求の範囲」が独特の書式で記述されていることから、調査に専門的な知識や経験が必要とされているためである。このままでは、特許調査に多くのコストがかかってしまう。そこで、「特許請求の範囲」は、常に名詞句の修飾を行っていることを利用して、ユーザにとって必要な名詞句に関する情報のみを効率的に提示し、ユーザの調査負担を減らすことを目的とした、ソフトウェア開発を提案する。

本提案では、まず、「特許請求の範囲」を名詞句と修飾語句単位に請求項にこだわらず分割し、名詞句をノード、修飾語句をリンクとした階層構造で表現する。階層構造は、同じ意味の名詞句を保持したノードを保持しないようにし、そして全体や一部というようにユーザが得たい情報に効率的にアクセスできるよう保持する。次に、通常が表示として、「特許請求の範囲」全文を名詞句ノードと修飾語リンクを用いて可視化する。そして、ある名詞句の下の階層の関係や、その名詞句を参照している他の請求項の下の階層の関係を可視化することで、特定の名詞句を中心とした修飾関係を再現する。

以上の機能から、ある部品がどういう状態の時に何をするものなのかといった情報を解りやすくユーザに提示するソフトウェアを作成する。

## 8. 採択理由

膨大な特許情報が溜ってきている現在、結構役に立つシステムができそうな提案である。文章を扱う課題にはいつも自然言語処理の難しい問題が絡むのだが、特許の文書という、人間にはわかりにくいのが、逆にコンピュータには処理しやすい文章を対象として、それを人間にわかりやすく、構造を可視化して見せるという提案である。なるほどどうまいところを突いている。

このシステムの発想の元となった既存研究があり、それをベースに増満君は卒論である程度の修行を積んでおり、さらに彼自身が弁理士志望なので実現性は高いというか、インセンティブが高い。プログラミング能力や、システム設計の詰めなど、まだ甘いところがあると思うが、課題が具体的であり、自分のメシのタネでもあるので、これはガッツをもって当たれば、プロジェクト期間中になんとでもなる。実際、増満君は、最初フリーソフトとして提供してユーザを獲得、その後、ビジネス化に進むというプランもっている。ところで、増満君のプレゼンを聞いて、彼の所属研究室ではなかなかいい薫陶が行なわれていると感心した。世の中にはユニークな研究室があるものだ。

## 9. 開発目標

「特許請求の範囲」の記述方法は、名詞句の定義を請求項 1、名詞句の条件を請求項 2、名詞句の動作を請求項 3 のように一つの要素に対して複数の項で分割して記述されている場合がある。これでは、見落としや誤解を生んでしまっても仕方がない構造と言える。このような場合は、その要素がどのような構成なのか知るために名詞句の関係をツリー構造で保持し、葉の要素を確認することで正確に把握することが可能であると考えた。この考え方を基本とし既存技術を発展させ、特許請求項解読支援を行うシステムを作成する。

## 10. 進捗概要

実施計画内容との相違点は、プラグインシステムを実装できなかったことである。

## 11. 成果

本成果は、特許文書を入力として特許請求項の可読性向上を目指した可視化を行った、主要な名詞句の抽出や省略を行うことで、発明を構成する名詞句の理解度の向上を行うことができた。既存システムに比べ、より深い階層での解析と請求項をまたいだ係り受けを解析しネットワークを形成することができたと考える。

## 12. プロジェクト評価

諸般の事情で契約が遅れ、開発に集中的に専念できない期間があったが、増満君はそれをなんとか跳ね返して、一定の成果が出るまでこぎ着けた。なかなかの精神力である。

採択理由にも書いたように、彼自身が弁理士志望なので、このプロジェクトに対するモチベーションが非常に高い。プロジェクト期間中に IT に強いとある弁理士事務所に丁稚奉公に入り、いろいろな研鑽を積み、システムの設計に役立てたようだ。さらにそこへの就職も内定し、そこでひょっとして新たなシステムの開発を行なえるかもしれないとのことであった。間接的だが、未踏ユースが新しい特許関連ソフトの市販または運用に結び付く事例になるかもしれない。

このプロジェクトのポイントは二つある。一つは特許文書特有の日本語の構文解析である。定形文書なので、やさしいかと思いきや、既存の多くの構文解析ソフトではうまくいかないものが多かったようだ。そのため、独自にいろいろなルール（増満君の言葉では IF/THEN ルール）を付け加えた。しかし、付け加えすぎると矛盾したルールも出てきてしまい、もぐら叩き状態になる難しさがある。特許文書とはいえ、自然言語で書かれた文書なので、どこかで本質的な壁に当たるものようである（慣れた専門

家にもとても読みにくい特許文書があるらしい)。現状では特許文書の 70%以上がまず問題なく解析できるとのことである。それでも、見なければいけない特許文書の的を早く絞るには十分な効果があるだろう。

もう一つのポイントは、解析結果の可視化である。可視化はユーザの要求が多様であるから、設計ではそれにどう応えるかが難しい。現時点では、可視化の基本素材がようやく揃ったという状況であるが、そのままでも一般技術者が特許文書を見るにはある程度役に立ちそうなレベルになっている。これをもっと突き詰める作業が必要だろう。ただし、カスタム化のメニューを増やしすぎて、却って使いにくくすることのないようにするスマートな設計哲学をもたなければならない。

このようなシステムが一般の技術者に使えるようになることは、知財立国を謳っている日本の技術戦略にとって重要なことである。散発的ではあるが、この種のシステムが市販されている、あるいは市販されそうな状況である。このタイミングで技術的な可能性を示すことは重要である。このプロジェクトにはそのような意義がある。

### 13. 今後の課題

動作検証の際に明らかとなった IF/THEN ルール実装のミスと、より細かい特徴を追跡し、長文化した修飾語句をうまく提示する方法を考察する。

また、解析のアルゴリズムはコンポーネント化しているため、Web 上でサービスとして展開したいと考える。今後は、特許調査のデファクトスタンダードを目指すべく公開に関する交渉と、実用性向上のためのシステム作成を行う。