

判決文における勝敗判別のための情報抽出

Extraction of winning or lost lawsuit information for legal precedents

伊藤 星子*1 松井 藤五郎*2 大和田 勇人*2
Seiko Ito Tohgoroh Matsui Hayato Ohwada

*1東京理科大学大学院 理工学研究科 経営工学専攻

Department of Industrial Administration, Graduate School of Science and Technology, Tokyo University of Science

*2東京理科大学 理工学部 経営工学科

Department of Industrial Administration, Faculty of Science and Technology, Tokyo University of Science

The trial of the United States works out with the case law. There is a tool of the retrieval of judicial precedents such as Lexis and Findlaw as a means to look for this useful judicial precedent, and examining the judicial precedent becomes possible by inputting the key word etc. The lawyers can easily find the case by inputting keywords with the tools, but the beginners can hardly find them.

We focus attention on the victory or defeat of a comprehensible judicial precedent as the means to retrieve an advantageous judicial precedent i for beginners.

In this paper, we propose the method consists of extraction module and analysis module.

1. はじめに

アメリカの裁判は、判例法から成り立っている。判例法とは、同じ種類の事件に対する過去の裁判において、同じ趣旨の判決が集積し、規格化され、将来の事件に対して先例として拘束力を有するようになったものである。裁判における判例の存在は重要なものであり、自分が裁判を抱えたとき、いかに有利な判例を探ることができるかが、結果を左右するポイントとなっている。この有用な判例を探す手段として、Lexis, Findlawなどの判例検索ツールがあり、キーワード等を入力することにより、判例を調べることが可能となっている。これらの判例検索ツールは、法律家にとっては詳細なキーワードを入力することにより的確な判例文書を探すことが可能であるが、一方で初学者にとってはどのようなキーワードを入力すれば良いのかわからず、的確な判例を検索することが困難であることが現状である。

そこで本研究では、裁判を抱えた場合より有利な判例を検索する手段として、初学者にもわかりやすい判例の勝敗を明確にすることに着目した。この判決の勝敗は従来の検索ツールでは、判例文書を読まなければ理解することが困難であるため、本研究では、判例文書から勝敗に関連する情報の抽出方法を提案する。

2. 判例文からの情報抽出システムの構築

本節では、本研究で提案する情報抽出システムの概要を示す。本研究では、判例文が一定の形式で記述されている点を利用して抽出方法を提案する。本システムの構成図を図1に示す。本システムは、抽出モジュールと解析モジュールから構成されている。抽出モジュールは、判例文書から勝敗に関連する情報を抽出し、データベースに格納する。また、抽出情報を判例文書1つに対して1つの抽出結果を出力する。解析モジュールは、抽出モジュールが判例文書から抽出した情報の解析を行う。ここでは、いくつかの判例文書について勝訴、敗訴といったカテゴリーごとに判例の分類を行う。

以下に判例文書の構成と、各モジュールの説明を行う。

2.1 判例の構成

判例文は、図2のように一定の形式で構成されている。

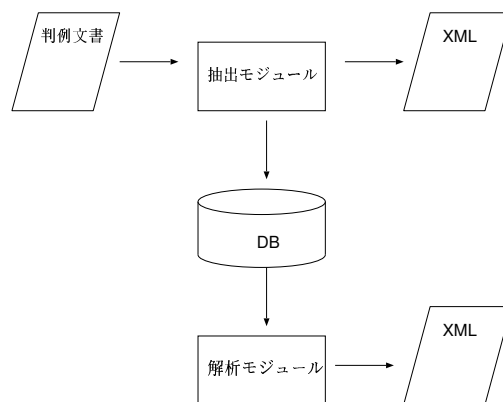


図 1: システム構成図

まず、行われた裁判所名が書かれており、順に裁判に関連する人物、一番の結果とともに今回の裁判の結論にいたるまでの理由が書かれ、最後に結論が書かれている。理由の裏付けとして過去の判例が挙げられている。また、結論だけではなく行われた裁判所、前回と今回の結論を出した裁判官名、当事者の弁護士名等の情報も書かれている。判例文は、一定のフォーマットで書かれているケースが多い。これは、文書の構成だけではなく、使われる表現についても一定の形式で記述される傾向がある。

2.2 抽出モジュール

抽出モジュールでは、判例文書から勝敗に関連する情報の抽出を行う。

2.1で述べたように、判例文は一定の形式で記述されていることから使われる表現や記述ルールを発見し、抽出したい情報の前後にある文字列を検索することにより、情報を抽出できると考えられる。抽出する情報は判決結果情報を構成する、裁判所名、裁判官名、事件名、原告名、被告、弁護士名、訴訟事件簿番号、引用判例と引用判例を使用した主張、判決結果とした。

例えば、引用判例の情報を抽出する場合は

1. 引用判例を用いた主張。

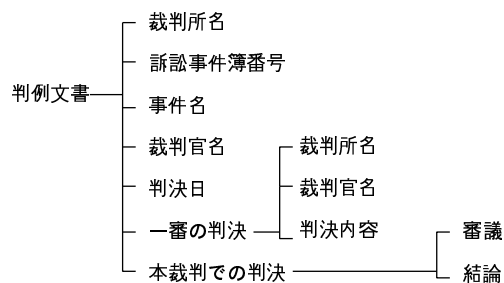


図 2: 判例文書の構成

2. 事件名 (A v.B),
3. 訴訟事件簿番号 (No.C-D).

の順に記述されているため、まず、判例名の v. で書かれている文字列について検索を行い、事件名を特定する。事件名を特定した後、その判例を用いた主張は事件名の前に書かれていることから、この文章を抽出する。

以上のようにして各項目ごとに抽出した情報は、その情報が何であるのか意味を保持させるため XML タグを付与する。XML タグは表 1 のように、抽出情報と対応する名称のタグを定義した。XML タグを用いることにより、抽出した情報に対応した名称を付けることができるので、データベースに格納した際に解析しやすいと考えられる。

以上のようにして、判例文書から抽出された情報は、XML タグを付与し 1 判例に対して 1 つのデータ文書を出力する。これらの出力結果は、二次利用性を考慮してデータベースに格納する。

表 1: XML 要素名と役割

XML 要素名	役割
court	裁判の行われる裁判所の名称
judge	裁判官名
plaintiff	原告の名称
defendant	被告の名称
case_name	引用判例と出典
case_reason	引用判例を理由とした主張

2.3 解析モジュール

2.2 によりデータベース化した判例データは、つぎに要素ごとに解析を行う。

抽出モジュールによって抽出された判例情報は、裁判官名、原告名、被告名、弁護士名、引用判例名ごとに、以下のステップで出力する。また、各要素ごとに抽出された情報が複数あるものについては出現回数をカウントし頻度の多いものほど重要であると解釈してランキングを付け表示する。

1. 抽出した各要素ごとに分類する。
2. 分類した各要素について、抽出情報の出現頻度をカウントする。
3. カウント数の多いものから順に出力する。

例えば、ある判例について引用判例として複数の事件に引用されている場合は、その事件 1 つに出現した場合を 1 としてカウントを行っていく。このようにして、各要素ごとに出現頻度をカウントしランキング化し出力するようにする。

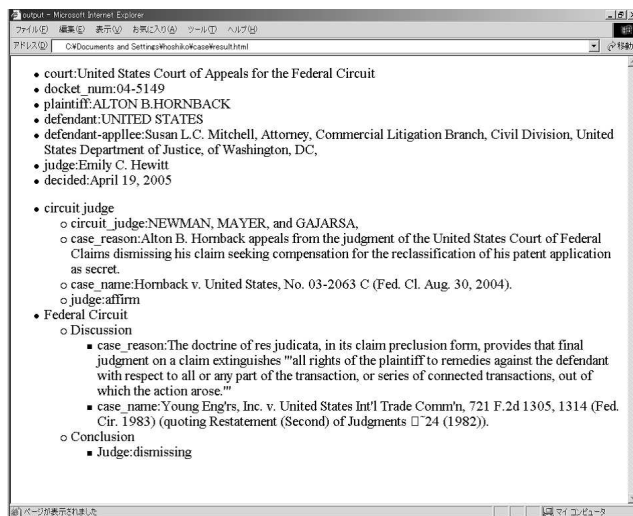


図 3: 抽出モジュールの出力結果

3. 実行例

本論文で提案した、判例文書からの情報抽出システムの有効性を示すため、控訴裁判所 (Appellate Court) で 2000 年 9 月から 2004 年 12 月までの patent における判例について、抽出モジュールで抽出を行った。

図 3 は、Alton B.Hornback v. United States(No.04-5149) の実行結果である。判例文書から、裁判所名、訴訟事件簿番号、被告名、原告名、1 番の判決結果、控訴審での審議内容および結果を抽出し、表示することができた。

本システムによって、判例文書から必要となる情報のみを出力することによって、判例文を読まなければ、勝敗について理解することができなかった点を必要情報だけ抽出し、表示することにより解消することができた。

4. おわりに

本研究では、従来の判例検索ツールでは、判例文書を読まなければわからなかった勝敗に着目しその判別を行うための情報を抽出するシステムを抽出モジュールと解析モジュールという形で提案した。

今後の課題として、抽出モジュール、解析モジュールともに提案段階であるため、完全なシステムとして動作できるようにし、この手法の有効性を示す。

参考文献

[Lexis] <http://www.lexis.com/>
 [The United States Court of Appeals for the Federal Circuit] <http://www.fedcir.gov/>
 [Bruninghaus 01] S. Bruninghaus and K. Ashley:Improving the Representation Of Legal Case Texts With Information Extraction Methods.,ICAIL-01 pp.42-51. ACM Press.
 [Bruninghaus 97] S. Bruninghaus and K. Ashley:Learning to Classify Case Opinions under Abstract Fact,ICAIL-97,pp. 123-131. Association for Computing Machinery: New York, NY.