

法的推論システム HELIC-II の判例ベース検索の拡張にむけて

石川 幹人 (松下電器産業株式会社*)
新田 克己 (電子技術総合研究所**)

HELIC-IIは、法令文と過去の判例とを利用し、新しい事件に対する法的判断例とその理由づけを生成するシステムであり、法律の実務や法学研究への計算機による支援を目指したものである。ところが、HELIC-II上での処理は、特有の形式での知識表現でもって行われるため、現在のHELIC-IIを法律分野の実用的な利用に供するには問題があった。毎年相当数が増加する判例データを活用するには、自然言語の原文レベルでの処理も可能なように、HELIC-IIの拡張がなされる必要がある。本論文では、知識表現に対応した部分原文を保持し、それに基づく照合を行う方策を提案する。

Toward the Improvement of Precedent Retrieval on Legal Reasoning System HELIC-II

Masato Ishikawa (Matsushita Electric Industrial Co.)
Katsumi Nitta (Electrotechnical Laboratory)

HELIC-II generates a legal judgment and its argument process for a given case based on the statutes and precedents. Although HELIC-II was designed to support the works on lawsuits, its special knowledge representation format has prevented the lawyers and legal researchers from utilizing the system. The precedent retrieval mechanism should be improved to interpret the original precedent text. In this paper, we propose the partial text matching method to extend the function of HELIC-II. The method is to hold the partial text which corresponds to each part of knowledge representation format, and to match the partial text with some part of given text.

1. はじめに

筆者らは、1991年に事例ベース推論技術に基づき、労働災害認定に係わる訴訟事件を例題として、過去の判例に基づく法的推論システムを開発した[1]。さらにそのシステムには、法令条文に関するルールベース推論処理部が加わり、刑法の訴訟事件を例題として、法的推論システムHELIC-IIの形態をなした[2,3]。その後、さらに価値判断、論争処理部が拡張されて今日に至っている[4]。こうしたHELIC-IIに関する一連のソフトウェアは、無償で一般に公開されている[5]。

HELIC-IIは、法令文と過去の判例とを利用した論証生成のシミュレーションを行うシステムであり、決して計算機による判決の自動生成を狙ったものではない。あくまで法律の実務や法学研究の上で、関連判例における判断や論証を調査したり、関連法令文における解釈の対立点を調査するに際しての、計算機による支援を目指したものである。しかし、HELIC-II上での処理は、特有の形式での知識表現でもって判例や法令文が記述されていることを必要としているため、システム開発者が知識表現を書き下す作業が膨大となっている。法令文の場合は長さも短く、数も限られていることから、知識表現の作業もそれほどではないが、判例の場合は長大で、毎年相当数が増加するため、知識表現の作業には自ずと限界がある。現在のHELIC-IIでは、こうした限界に対する対応策を確立することが、

* 明治大学法学部講師を兼務

** 東京工業大学大学院総合理工学研究科教授を兼務

法律分野での実用的な利用に向けての最大の課題となっている。

本論文では、HELIC-IIにおける判例ベースの検索利用を、手に入る判例データがすぐに利用できるよう、拡張する方策について議論する。まず、第2章で判例データについて概観し、第3章でHELIC-IIについて、とくに事例ベース推論部分を中心に概説する。続く第4章で、HELIC-IIの判例データ検索・処理に関わる問題点を指摘し、改善策を示す。

2. 判例データ

わが国における判例データを備えた最初の情報システムは、日本電気が1968年に開発したJUSTICEである[6]。JUSTICEは、最高裁判所の判例をカタカナのキーワードでもって検索する機能を持ち、その後の判例検索のパイロットシステムとなった。1987年には、梅本吉彦専修大学教授を研究代表者とするグループにより、判例の本文全文を格納し、それを検索対象とした交通災害判例検索システムが開発された[7]。同システムでは、日立製作所のHITAC-Mシリーズ上に交通災害に関する判例を格納し、検索ソフトウェアORIONを使って検索機能を実現した。

また、官公庁では、1980年に総務庁（当時の行政管理庁）が行政官庁の共同利用サービスの一貫として、法令判例検索システムを開発した[8]。判例データについては、現憲法下の判例のうち、判例集に記載の全件を入力しているが、裁判要旨のみで本文は入力されていない。通産省、建設省、農水省等でも独自に、各省に関連する法令判例データを整備し、それらに関しても同様の検索サービスを確立している。しかし、これらの官公庁におけるサービスは、省庁外には開放されていない。

一般利用者向けの最初の実用システムは、1984年に東洋情報システムが開発したJUPITERである。JUPITERは最近まで市販されていたが、CD-ROMによる情報の廉価提供の流れに圧されたためか、現在では廃止されている。現在のところ、市販で手に入る電子化された判例データは、以下に掲げたものである。

(a) 判例体系（第一法規出版）

CD-ROM（Windows版）9枚に、現/旧憲法下の判例約12万5千件を収録。うち約9万5千件については本文（全文：主文と事実と理由）も収録。文中から切り出した語句に基づくインデックスを対象としたキーワード検索。年2回更新。法編毎の分割購入可。

(b) 判例MASTER（新日本法規出版）

CD-ROM（Mac版/Windows版）1枚に、現憲法下の判例9万2千余件を収録。うち代表的な約8千件については本文（主文と理由のみ）も収録。6万語の辞書に基づくインデックスを対象としたキーワード検索。年2回更新。NIFTY-serveのホームパーティ「判例の達人」で最新情報を随時提供。

(c) リーガルベース（日本法律情報センター）

CD-ROM（Windows版/Mac版）1枚に、現憲法下の判例約9万2千余件を収録。全件について本文（必要とされる重要部分のみ）も収録。完全一致全文検索。ハードディスク等へのダウンロードに制限なし。年2回更新。DVD対応も準備中。

(d) 行政判例CD-ROM（ぎょうせい）

CD-ROM（Windows版）1枚に、地方自治に関係する判例に限定して、現/旧憲法下の約1万件を収録。全件について本文（判決理由の一部のみ）も収録。類義語展開が可能な完全一致全文検索。年2回更新。

(e) 模範六法CD-ROM版（三省堂）

CD-ROM（電子ブック規格）1枚に、模範六法が収録されたもの。法令の各条文につき、関連する判例要旨が約8千件付記されている。電子ブック用検索ソフトを利用。数年毎に改版。

(f) LEX/DB（TKC）

(a)の判例体系と同様のサービス内容を、パソコン通信ネット等からオンラインで提供。データは毎週更新。検索代行サービスも有り。

以上に示した各々の判例*データの判例件数は、例えば特許庁が提供している出願公開特許の電子データ（平成6年分で年間35万件[9]）に比べ、極めて少ない。ところが、実際の民事・行政事件の裁判例数[10]は、平成6年分で年間240万件であり、平成2年の178万件から4年間で35%の伸びを示している。この裁判例数は、特許の出願数に比べ1桁多い。すなわち、わが国における判

* 拘束力をもつ判例は最高裁判例に限るなどの議論[13]があるが、本論文での判例は単に全裁判例を指す。

例の電子データは、特許明細書の電子データのように全件全文は完備されていないのである。一方、米国では、LEXIS、JURIS、WESTLAW等の判例データベースシステムが、1970年代から整備されてきており、分野別のライブラリ化や、オンラインによる検索が実現されている[11]。LEXISは、1790年以降の約600万件のデータが収録されており、現在でも週に1万件のペースで増加している。

わが国において、判例データの電子化が遅れている理由は、第1に裁判例の公開・電子化が進んでいないこと、第2に米国のように先例拘束主義ではなく、大陸系の成文法主義であるため、判例公開の要求が比較的低いことが挙げられよう。その結果、現在、判例集等により公刊されている判例は、高裁・最高裁判例を中心に年間2千件程度に過ぎない。先に挙げた各社の判例データは、各々、公刊された印刷文書に基づいて手作業で入力したものである（要旨、解説等の著作権がある部分は、別途、各社独自に作成している）。

しかし、最近では、社会における情報公開の要請が高まりつつあるうえ、裁判実務の情報化の兆しも現れている[12]。個人情報の保護が留意されながらも、多くの判例が電子的に公開されるのも、そう遠くはない将来のことであろう。また、わが国は成文法主義をとっているといえども、判例は抽象的法規範の具体化や、法文に規定のなかった一般法理の定立、あるいは社会の実態に即した法文の実質的修正などに積極的な役割を果たしている[13]。近年のように、情報化・グローバル化という言葉で言い表されるように社会の変化が激しくなってくると、判例の果たす役割も一段と大きくなってくる。そのため、判例に関する洩れない調査は、実務においても、法学研究においても重要な課題となり[14]、判例データの完備とその円滑な検索利用に対する要求は、以前にも増して大きくなるに違いない。

3. 法的推論システム HELIC-II

HELIC-IIは、図1のように2つの推論エンジンと法令文のルールベース、判例の事例ベース、そして概念辞書からなる。一般に判例文には、事件の概要、双方の主張、裁判官の理由づけと結論等が書かれているが、HELIC-IIでは、判例の情報を「状況記述」と「事例ルール」という知識表現形式で表す。「状況記述」は図2上のように、事件の概要を、オブジェクトやそれらの間の時間関係で記述したものである。オブジェクトは、人物、行為、権利等を、「概念を表す述語」と「オブジェクトの識別子」と「そのオブジェクトに関する性質を[属性名=属性値]の対のリスト」で表現したものである。属性値もオブジェクトとなる。システム内部では、この記述は、オブジェクトをノードとし、属性（あるいは時間関係）をリンクとした意味ネットワークとして展開される（図2下）。

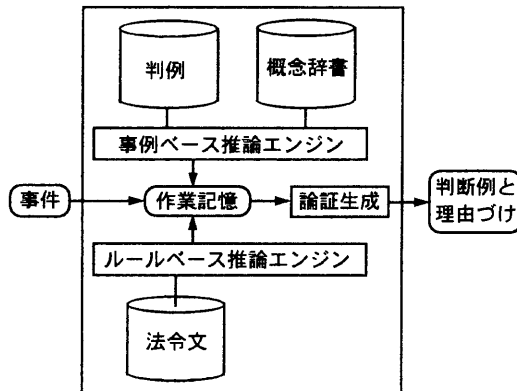


図1 HELIC-IIのシステム構成

case(丙事件, [meets(運転1, 事故1),
after(事故1, 状態1), ...]).
自然人(丙, [性=男, 年令=34]).
自然人(山田, [性=男, 年令=40]).
運転(運転1, [主体=丙, 対象=自動車1]).
交通事故(事故1, [主体=丙, 客体=山田]).

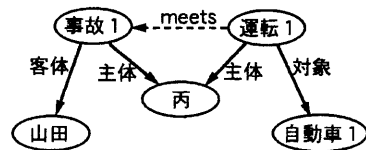


図2 状況記述

「事例ルール」は、当事者双方の主張や裁判官の理由づけを「～の条件に適合する → ～と帰結する」の形式のルールの集合で表現したものである（図3右）。「事例ルール」には、事実解釈ルールと法的判断ルールがあり、事実解釈ルールは、事実関係から法的概念を導くもので、具体的な概念を用いて記述され、条件部は類似性に基づいて適合される。それに対し、法的判断ルールは、ある法的概念から次の発展した法的概念を導くもので、抽象的な概念で記述され、変数を含み、条件部はより

完全な適合が要請される。法令文ルールの形式も法的判断ルールと同様であるが、法令文ルールの場合、類似性は考慮されずに条件適合が行われる。裁判での論理展開は、「状況記述」を開始点とする、事例ルールと法令文ルールの条件適合連鎖（推論）と捉えることができる[15]。

事例ベース推論エンジンの役割は、新しく検討すべき事件が与えられたときに、類似の判例中の論理展開を利用して、「公共の福祉に反する」などの抽象的な法的概念に至る主張を構成し、列挙することにある。これは、新しい事件の状況記述を手がかりにして、類似事例検索と類似論理構築の2つの段階を経て行われる。類似事例検索では、新しい事件の状況記述と、判例ベースにある各判例の状況記述を比較することで、類似する判例を粗く選定する。類似性の判定は、状況記述を構成する意味ネットワークのリンクを相互に照合し、照合がとれた数によって行う。各リンクの照合は、同一名のリンクの両端のオブジェクトについて概念辞書を参照し、そこに書かれている概念体系の階層構造において、許容される距離内で上位概念を共有する場合は、照合と判定する。一般用語の概念辞書は電子化辞書[16]様の形式を使用しており、それに法律用語の概念体系を付加している。

続いて類似論理構築では、選定された各判例の事例ルールにつき、新しい事件の状況記述と類似性に基づく条件適合を行い、法的概念の仮設を生成する。生成された法的概念は、また他の事例ルールにより、次の法的概念が仮設される。類似性に基づく条件適合は、事例ルールの条件部に現れる記述と、新しい事件の状況記述とを比較し、リンク単位で照合したうえで条件部全体で適合を判定することでなされる。リンク単位の照合は、類似事例検索と同様に概念辞書を参照して行う。条件部全体の適合判定は、条件部の記述にあらかじめ付与してある重要度の重みに依存する。重みのうちexactが付されているリンクは、そのリンクが照合しない限り条件適合はなされない。重みには他にimportant (imp)とtrivialがあり、それが付されたリンクについては概念辞書中の階層距離に相対的な重みづけをしたうえで、条件部全体の類似度を積和計算する。その類似度が設定した値を越えたならば条件適合がなされる。結果として、あるノードにおいて概念辞書中で上位概念を共有してなくとも、意味ネットワーク上で周囲のリンクが照合していれば、類似性に基づく条件適合がなされる場合も出てくる。

4. 判例ベース検索の拡張にむけて

HELIC-IIは、上に述べたように法令文と過去の判例とを利用し、新しい事件に対する法的判断例とその理由づけを生成するシステムであり、法律の実務や法学研究への計算機による支援を目指したものである。ところが、現在のHELIC-IIを法律分野の実用的な利用に供するには、次の解決すべき問題がある。第1に、極めて多くの判例件数にわたり、いちいち「状況記述」や「事例ルール」を書き下しては、システム開発者の作業が膨大になってしまう（判例記述の問題）。第2に、増加する判例に応じて必要となる概念辞書の整備・更新の手間が煩わしい（辞書更新の問題）ので、自動更新[17]の要求がある。第3に、検討すべき新しい事件を入力するにあたり、システムの利用者がその事件の「状況記述」を特有の形式で表現するのは無理がある（新事件入力の問題）。第4に、結果として得られた理由づけを出力するにあたり、「事例ルール」の条件適合連鎖をそのままの形式で表示しても、システムの利用者は容易には理解できない（理由づけ出力の問題）。これらの問題に対応するには、自然言語の原文レベルでの処理を拡張する必要がある。

本論文で提案する解決策*では、各知識表現形式に添えてそれが由来する部分原文を保持し、その部分原文を処理すべき判例原文に対して照合し、改訂した知識表現形式を、対応する知識表現形式に基づいて生成する。詳述すれば、以下の3段階で対処する。第1段階では、代表的な判例について、これまでと同様にシステム開発者が手作業で事例ルールを作成する（図3）。そのうえで、対応する原文もデータ蓄積する。例えば、図3右の事例ルールに対して、図3左の部分原文データを対応づけて格納する。状況記述の作成も同様に手作業で行い、対応部分原文とともに格納する。ただし、事例ルールの条件部に出現しないような状況記述は必要ないので、知識表現形式に書き下さない。

第2段階では、代表的判例以外の判例の知識表現形式を、判例原文から自動生成する。そもそも法律文における語句の使用形態や言い回しは、自然言語一般に比べ限られており、とくに法令文ではそれを制限言語として捉えた言語解析研究がなされている[18,19]。判例においても、事実関係の記載等は一般的な非定型の文体であるものの、主張や理由づけの記載部分は比較的定型性が高い。そこで、あらかじめ手作業で切り出した部分原文と対応事例ルールとを手掛かりにすれば、改訂事例ルー

* 遺伝子情報処理の分野に、配列データから遺伝子の構造・機能を予測するに際し、すでに構造・機能が知られている配列データに加えて、構造・機能が知られていなくとも、それと類似の配列データを集めることによりデータ量を増やし、予測率を飛躍的に上げる方策が知られているが、本解決策はその類推である。

「栗原は前示のように被告人等から暴行を受け、なおも被告人等から危害を受けることを恐れ、これを避けるため、救を求めながら逃走したが、被告人等から包囲体制をとって追跡された結果、逃げ場を失い、やむなく江戸川に飛び込み溺死したものであるから、被告人等の暴行と栗原の死亡との間に因果関係があるものと認めるを相当とするのである（東京高判昭32年・高判10-3-313）」



```
rule([meets(暴行1, 逃走1), equal(逃走1, 追跡1),
      meets(逃走1, 溺死1)],
      暴行(暴行1, [主体=被告人, 客体=栗原]/exact),
      畏怖(恐れ1, [主体=栗原, 客体=被告人]/imp),
      避難(逃走1, [主体=栗原, 対象=江戸川]/exact),
      追跡(追跡1, [主体=被告人, 客体=栗原]/imp),
      死亡(溺死1, [主体=栗原]/exact),
      --> 因果関係(関係1, [原因=暴行1, 結果=溺死1])).
```

図3 事例ルールの作成

「前述のとおり同人は被告人に追われて止むを得ずこのような海中に転落する危険性の極めて高い不安定な場所に逃れたものである以上、同人が手を離れた直接の原因が前述のいずれであるとしても、なお同人の転落と被告人の暴行行為（追跡行為を含む）との間には法律上の因果関係があるものといわなければならない（広島地判昭47年・刑月4-12-2000）」



```
rule([meets(暴行1, 逃れ1), equal(逃れ1, 追う1),
      meets(逃れ1, 転落1)],
      暴行(暴行1, [主体=被告人, 客体=同人]/exact),
      避難(逃れ1, [主体=同人,
                  対象=不安定な場所]/exact),
      追跡(追う1, [主体=被告人, 客体=同人]/imp),
      死亡(転落1, [主体=同人]/exact),
      --> 因果関係(関係1, [原因=暴行1, 結果=転落1])).
```

「証拠によれば、被害者は被告人らから足蹴りされる等の暴行に耐えかね隙をみて逃げ出し斜面に岩石の露出している池に落ち込んだもので、その後頭部上部正中を打撲したとすれば被告人らの加えた暴行と前掲打撲傷に基づく死亡との間には刑法上因果関係があると認めるのが相当である（最三小決昭59年・刑集38-8-2803）」



```
rule([meets(暴行1, 逃げ1),
      meets(逃げ1, 打撲1)],
      暴行(暴行1, [主体=被告人, 客体=被害者]/exact),
      避難(逃げ1, [主体=被害者, 対象=池]/exact),
      死亡(打撲1, [主体=被害者]/exact),
      --> 因果関係(関係1, [原因=暴行1, 結果=打撲1])).
```

図4 事例ルールの自動生成

ルを生成する程度の言語処理は可能と思われる。例えば、処理すべき判例原文に図4左のような記載があった場合、図3左の部分原文と照合し、図4右の事例ルールが生成できよう。照合は文法構造における語句の対応づけ*と、前章と同様の類似性判定（重みづけと概念辞書による）によって行う。例えば、「因果関係」「暴行」は両原文で合致するが、図3左の「逃走」は、図4左では「逃れ」「逃げ」となっている。概念辞書を参照して該当部分の照合を行い、図3右の「逃走」の部分を入れ換える。また、図4上では因果関係の原因と結果の順番が入れ替わっているが、構文の並列関係を調べることで照合がとれる。手作業で作成した事例ルールと全く異なる事例ルールは生成できないものの、この方法で条件部に変更を加えた改訂事例ルールが生成できる。事例ルールの生成が行われたならば、それに関連する状況記述も、格納してある状況記述と、それに対応する部分原文とを手掛かりにして同様に生成が可能であろう。ただし、状況記述に関する記載は定型性が低いので、照合における類似性判定の許容度合を大きく下げることが必要である。

第2段階までで、多数の判例にわたって知識表現の形式になったデータが揃う（判例記述問題の解決）。似たような事例ルールが多く蓄積されて同時に照合が発生するという問題が起こりかねないが、似た事例ルールをまとめあげて統計的な主張の強さに還元させれば、かえって有効に活用できるだろう。また、改訂事例ルールと、元の事例ルールとを比較することで、概念辞書の更新も行える。例えば、図4右の「転落」が概念辞書中になかったとしても、図3右の「溺死」と同様に「死亡」の低位概念として概念辞書を更新すれば、妥当性に問題がある場合もあるので考慮が必要だが、その後の類似性判定に活用できる（辞書更新問題の解決）。

第3段階では、新しい事件の状況記述を生成する。新事件についてはシステムの使用者が新規に入力するので、散文調になることもあり、文法的な解析の可能性はあまり期待できない。だから、判

*すでに、企業提携とマイクロエレクトロニクス分野の新聞記事を題材にした、この種のテンプレート照合による情報抽出の研究が、コンテスト形式で行われている[20]。

例データに蓄えられている状況記述と語句レベルの粗い照合をして、新事件の状況記述を生成するとよいだろう。新事件の事実関係が網羅的に入力されることも、同様に期待できないので、比較的多くの状況記述が照合する事例ルールの条件部に関しては、残りの照合しない状況記述をも新事件の状況記述として仮設する等の、特別な対処も必要であろう。その結果、HELIC-IIの推論処理が、新事件についての処理というよりは、新事件と類似の判例での論理展開を列挙する処理に近くなってしまふ*が、それでも法律分野での利用という目的には十分かなうであろう（新事件入力問題の解決）。さらに、推論結果の理由づけにあたる論理展開の列挙に際しては、事例ルールの条件適合連鎖をそのまま表示するのではなく、対応する部分原文を、新事件に照合した語句を埋め込んだうえで連鎖状に出力すると自然言語表示になるため、システム使用者の理解が容易となろう（理由づけ出力問題の解決）。

5. おわりに

HELIC-IIを法律分野の実用的な利用に供するにあたって、解決すべき問題点について議論し、可能性のある解決策を提案した。今後の課題は、解決策の実装と評価である。それに伴って、代表的判例の選び方、部分原文照合の高速な処理法、照合における類似性判定の許容度合、改訂事例ルールの複合化（条件部の拡張）や分離（推論段階の挿入）法等の、具体的事項の検討が必要となってくる。

参考文献

- [1] Hoshida,M., Nitta,K. and Ishikawa,M.: Experimental System of Parallel Legal Reasoning using Precedents, Proc. Joint American-Japanese Workshop on Parallel Knowledge Systems and Logic Programming, UC Davis Research Report CSE-91-27, pp.M16-23 (1991).
- [2] 新田克己：法的推論システム HELIC-II, 人工知能学会誌, Vol.7, pp.603-607 (1992).
- [3] 大嶽能久, 新田克己, 前田茂, 小野昌之, 大崎宏, 坂根清和：法的推論システム HELIC-II, 情報処理学会論文誌, Vol.35, pp.986-996 (1994).
- [4] Nitta,K., Shibasaki,M., Sakata,T., Yamaji,T., Xianchang,W., Ohsaki,H., Tojo,S. and Kokubo,I.: New HELIC-II: a Software Tool for Legal Reasoning, Proc. International Conference on Artificial Intelligence and Law, pp.287-296 (1995).
- [5] ICOT Free Software (<http://www.icot.or.jp/>)
- [6] 堀部政男, 永田眞三郎：情報ネットワーク時代の法学入門, 三省堂 (1989).
- [7] 梅本吉彦：法情報検索の現段階, ジュリスト増刊「ネットワーク社会と法」, pp.130-135 (1988).
- [8] 高石義一：法律情報検索の現状と課題, にじゅういち出版, pp.82-144 (1985).
- [9] 特許庁総務部総務課「特許庁広報」
- [10] 最高裁判所事務総局総務局統計課「司法統計年報（民事・行政編）」
- [11] 田島裕：法律情報のオンライン検索, 丸善 (1992).
- [12] 夏井高人：裁判実務とコンピュータ, 日本評論社 (1993).
- [13] 中野次雄：判例とその読み方, 有斐閣 (1986).
- [14] 良永和隆：法情報検索の意義と方法, 専修大学出版局「法情報学要論」, pp.33-61 (1991).
- [15] Branting,K.: Representing and Reusing Explanations of Legal Precedents, Proc. International Conference on Artificial Intelligence and Law, pp.103-110 (1989).
- [16] EDR電子化辞書使用説明書, 日本電子化辞書研究所 (1993).
- [17] 樽松理樹, 山口高平：事例に基づく推論とモデル推論の統合に基づく知識獲得支援システム (1) 法律解釈知識の獲得, 人工知能学会誌, Vol.11, pp.585-592 (1996).
- [18] 野村浩郷：法律文制限言語モデルに基づく法律文の計算機処理, 吉野一編「法律エキスパートシステムの開発研究」, 文部省科学研究費研究成果報告, pp.229-252 (1995).
- [19] 川添一郎, 牧隆史, 田中規久雄：法律条文の標準構造(2), 情報処理学会自然言語処理研究会 107-13, pp.97-104 (1995).
- [20] MUC-5: Proc. Fifth Message Understanding Conference, Morgan Kaufmann Publishers (1993).
- [21] Salton,G.: The Vector Space Model, Automatic Text Processing, Addison-Wesley Publishing, pp.312-325 (1989).

*むしろ、新事件の状況記述生成は一切あきらめて、新事件の原文と判例ベースの部分原文とをベクトル空間法 [21]等で照合し、照合した部分原文に対応する知識表現を起点として、推論を開始させる方が良いかもしれない。