

司法試験の問題分類と法律文書における人物関係の構造化辞書

Problem classification of legal bar exam and structured dictionary design for human relationships in legal documents

清田 直希*¹
NAOKI KIYOTA

狩野 芳伸*¹
YOSHINOBU KANO

静岡大学情報学部*¹

Abstract: Legal document processing is one of promising applications of natural language processing. However, legal sentences are very complicated and require various difficult analysis techniques. We focus on the competition workshop COLIEE (Competition for Legal Information Extraction and Entailment) that aims to automatically answer the legal bar exam. Firstly, we manually classify the past legal bar exam data distributed in COLIEE by required analysis techniques such as conditional sentence extraction and negative form interpretation. Secondly, we focus on person relationships among our classifications, propose a dictionary design for legal document processing based on verbs and their arguments with roles of persons.

1. はじめに

近年の技術革新と共に注目を集めている分野が自然言語処理である。計算機が、人間が日常的に利用している自然言語を処理するために様々な研究がおこなわれている。

その中でも、法律文書の処理に関しては、一般的な文書の言語処理に加えて、専門性の高い言語が多く用いられている。そのうえ、文章中の事象を客観的に評価する必要があるため、難易度は高いが自然言語処理の分野において期待が高まっている。本研究では、毎年開催されている我が国の司法試験自動解答を行うコンテスト型ワークショップ COLIEE (Competition for Legal Information Extraction and Entailment)タスクシリーズ [1][2][3][4]で配布されたデータを対象とし、法律文書処理の基盤リソース構築を狙う。

法律文書においては先述の通り、専門性の高い言語に加え、構文、意味役割、言い換え表現など高度かつ多様な解析技術を必要とする複合的な要素がある。中でも、人物関係は司法試験を解くうえで重要な問題分類の一つであり、こうした処理には用語の構造化された整理が必要であると考えられる。そこで、人物関係に着目した構造化辞書を提案する。

2. 司法試験の自動解答タスク COLIEE

COLIEE は、COLIEE 2014、COLIEE 2015 [1]、COLIEE 2016[2]と人工知能学会国際シンポジウム(JSAI-isAI)のひとつである JURISIN(法情報学ワークショップ)において開催してきた。COLIEE 2017[3]は法と人工知能に関するトップカンファレンスである ICAIL (International Conference of Artificial Intelligence and Law)において開催された。COLIEE 2018 ではカナダ法を用いたタスクが追加されたが、本稿では我が国司法試験を用いたタスクのみを対象とする。

司法試験は弁護士や裁判官、検察官となろうとする者を対象に行われる試験である。問題形式は択一問題を含む短答式と論文式に分けられて行われる。試験の科目数については、短答式が憲法、民法、刑法の 3 科目であり、論文式が公法系科目(憲法及び行政法に関する分野)、民事系科目(民法、商法及び民事訴訟法に関する分野)、刑事系科目(刑法及び刑事訴訟法に関する分野)、選択科目(知的財産法、労働法などの中から 1 科目選択する)の 4 科目である。

COLIEE ではこのうち民法短答式試験を対象としている。多択式を二択にブレイクダウンするとともに、解答に関する条文を検索する情報検索タスク(Information Retrieval)と、二択を

Yes/No で答える含意関係タスク(Entailment)を用意した。また知識源として民法の全条分を配布し、問題文共々原文の日本語に加え、英語に翻訳したものを配布している。

COLIEE は与えられた問題文に対し、民法条文から関連する条文を返す情報抽出タスク(Phase 1)、問題文と与えられた関連条文について二値の安定をする含意関係タスク(Phase 2)、関連条文を与えずにこの二つを同時に行う質問タスク(Phase 3)の三つのサブタスクからなる。COLIEE 2017 においては、時期の問題から Phase 2 を実施せず、Phase 1 を Task 1、Phase 3 を Task 2 と呼称して実施した。以下で各タスクを定式化する。

情報抽出タスクは、与えられた問題文 Q を回答するのに関連する民法条文のサブセット S1, S2, ..., Sn を民法全体から抽出するタスクとなる。解答すべき条文は複数のこともある。

含意関係タスクは、与えられた問題文 Q を、同時に与えられた民法条文のサブセット S1, S2, ..., Sn が含意するか、すなわち Entails(S1, S2, ..., Sn, Q)または Entails(S1, S2, ..., Sn, not Q)を回答するタスクである。

質問応答タスクは与えられた問題文 Q について、民法全体を用いて Yes あるいは No を回答するタスクである。

3. 司法試験の要求要素技術別問題分類

司法試験を解くうえで、背後にある構文構造、人物の役割と関係性、論理、抽象性など多様で複雑な構造を的確に処理をする必要がある。そこで平成 29 年度のテストデータを対象に、解答に必要な要素を設定した基準を基に手作業で分類をした。表 1 は、解答に必要な要素を手作業で分類した結果と、COLIEE2018 参加チームの要素技術別の正解数及び正答数をまとめた表である。分類の基準は表 2 に記した。この分類は、要素技術ごとに排他的ではないため、ある問いに対して複数の要素技術を要することもある。以下では、分類した要素技術(タスク)のうち、テストデータに頻出なものについて簡単な説明を述べる。

3.1 条件文抽出

このテストデータにおいて、最も多くカウントされた要素技術は、条件文抽出である。これは問題文・条文のそれぞれの条件文を調べることで正誤判定ができる可能性がある問題である。例えば以下のような問いである。

第六十条 相続財産に関しては、相続人が確定した時、管理人が選任された時又は破産手続開始の決定があった時から六箇月を経過するまでの間は、時効は、完成しない。

	問題数	UA[6]	正答率	UE	正答率	KIS_mo3[7]	正答率	KIS_Frame[7]	正答率
条件文抽出	31	19	0.61	9	0.29	13	0.42	16	0.52
人物役割付与	27	16	0.59	12	0.44	14	0.52	13	0.48
人物関係抽出	26	14	0.54	11	0.42	13	0.50	10	0.38
形態素解析	25	18	0.72	16	0.64	20	0.80	16	0.64
照応解析	20	13	0.65	5	0.25	12	0.60	13	0.65
曖昧性解消	17	9	0.53	7	0.41	8	0.47	9	0.53
意味役割付与	15	8	0.53	6	0.40	9	0.60	6	0.40
動詞言い換え	13	7	0.54	6	0.46	7	0.54	4	0.31
一般辞書	9	5	0.56	5	0.56	4	0.44	4	0.44
述語項解析	9	8	0.89	3	0.33	5	0.56	5	0.56
否定形解釈	7	5	0.71	3	0.43	5	0.71	7	1.00
法律系用語	7	4	0.57	2	0.29	2	0.29	3	0.43
含意関係	5	2	0.40	2	0.40	1	0.20	2	0.40
係り受け	5	3	0.60	1	0.20	2	0.40	4	0.80
条文検索	5	3	0.60	2	0.40	3	0.60	4	0.80
言い換え	5	2	0.40	4	0.80	3	0.60	3	0.60
箇条書き	3	1	0.33	2	0.67	1	0.33	2	0.67

表 1. 平成 29 年度のテストデータの問題の要素技術別分類と COLIEE 参加チームの正答数及び正答率

問い: 相続財産に関しては、相続財産管理人が選任された場合でも、相続人が確定するまでの間は、時効は完成しない。下線部が文章中の条件と考えられる。問いでは相続財産の時効不成立の条件が正誤の判断材料になる。前者は「相続人が確定し、管理人選任(または破産手続き決定)から 6 カ月が経過するまで」という条件下では時効が完成しないのに対し、後者は「管理人が選任されて、相続人が確定するまで」という条件になっている。このように 2 つの文は条件が異なっているため、答えは No となるのである。

3.2 形態素解析

また、形態素解析が特に必要となる問題も数多く見られた。形態素解析は自然言語処理において重要なタスクの一つであるため、特記をする必要性は感じられないかもしれない。ここでの分類基準は、係り受け(単語の修飾関係)や述語項解析が判断基準となる問いの他、条文がおおむねそのまま問いになっているものも含まれる。つまりは形態素解析さえあれば解ける問いも少なからず見受けられるということになる。こうした意味でも形態素解析という分類基準を追加した。

3.3 人物役割付与・人物関係抽出

人物役割付与は、問題文に登場する人物(例えば A など)が条文上のどの役割を担っているのかを明確にするタスクである。人物関係抽出は、売買契約や婚姻関係等の複数の人物間の立場やその関係性を示すために必要となるタスクである。これらについては次節で述べる。

分類名	分類規準
条件文抽出	問題文・条文のそれぞれの条件文を調べることで正誤が判断できる可能性がある問題
人物役割付与	問題文に出てきた人物(A など)がどの役割(未成年, 買主, 債務者 等)を担っているのかを明確にする必要がある問題
人物関係抽出	複数の人物間の立場や役割を明確にする必要がある問題
形態素解析	おおむね同じような文章である問題 又は文章を解析しやすくするために特定の格助詞などを入れるのに用いる
照応解析	文章の一部を省略している際に、それが何を指しているのかを明確にする際に用いる。「こそあど言葉」のみだけではなく「前文」「同項」なども対象となる
曖昧性解消	曖昧な表現が用いられている文を解析するのに用いる 例: H29-4-A 2 前項の規定による意思表示の無効は、善意の第三者に対抗することができない。
意味役割付与	問題文の中で用いられている名詞が動作主なのか・その動作の対象なのか、手段なのか等を特定する必要がある問題
動詞言い換え	動詞が言い換えられている問題
一般辞書	法律系の言語ではない一般生活で用いられるような事象が出てきている問題(例: 冷蔵庫、会社)
述語項解析	その問題の動作を明確にする必要がある問題。特に「～した」「～できる」などはこうした動詞を抜き出すのではなく、その行為を抜き出す必要がある
否定形解釈	否定形になっていて正誤判定をする問題 又は「～できない(不可能)」「～しなければならない(義務)」など「ない」が用いられていてもその意味が異なる問題
法律系用語	法律に関する言語が言い換えられている問題(例: 制限行為能力者→未成年など)
含意関係	条文から隠された意図を把握する必要がある問題 ピカソが『ゲルニカ』を発表した→ピカソは『ゲルニカ』の作者である
係り受け	主語と述語の関係、単語の修飾又は文の並列関係に着目する必要がある問題
条文検索	条文中に別の条文が指定されている問題
言い換え	動詞以外で言い換えがされている問題(例: 1→1)
箇条書き	条文で箇条書きが用いられている問題。

表 2. COLIEE 司法試験問題の要素技術別分類基準

3.4 照応解析

照応解析は、文章中で登場する「これ」や「あの」などの指示内容を明確にする解析である。前述の指示詞に加えて、「前章」や

「同項」など司法試験に特有の指示表現もあり、言語処理においてはその内容を正しく理解する必要があると考えられる。

3.5 曖昧性解消

曖昧性解消は、文章中に登場する意味的に曖昧な用語を適切に捉える必要がある問いにつけるタスクである。曖昧な用語とは以下のようなものである。

第四百七十五条 弁済をした者が弁済として他人の物を引き渡したときは、その弁済をした者は、更に有効な弁済をしなれば、その物を取り戻すことができない。

ここで述べられている「物」というのが曖昧な用語に該当する。このほかにも名詞であれば「事由」や「意思」、動詞であれば「する」や「ある」といったものが該当する。こうした曖昧性は、複数の対象あるいは幅広い語義を含み、文章理解が困難である。

3.6 意味役割付与

意味役割付与は、文章中の語彙に動作主や場所、対象といった意味的な関係を付与するタスク、SVM[5]やニューラルネットワーク[5]を用いた機械学習手法の提案がされている。

3.7 否定形解釈

否定形解釈は、条文あるいは問いの否定形を正しく解釈したうえで、正誤の判断材料にするタスクである。条文の内容を否定形に反転した司法試験の問題も見られる。

第六百七十五条 組合の債権者は、その債権の発生の時に組合員の損失分担の割合を知らなかったときは、各組合員に対して等しい割合でその権利を行使することができる。

問い: 組合の債権者は、各組合員に対して、その権利を行使することができない。

このように条文では権利を行使できる旨が書かれているが、問いでは権利の行使ができないことになっている。このように否定形にすることにより答えが偽になる問題も司法試験では見られ、それは文末の「ない」を認識できれば簡単に評価できる。ところが、次のような問いについてはそうもいかない。

第六百三十三条 報酬は、仕事の目的物の引渡しと同時に、支払わなければならない。ただし、物の引渡しを要しないときは、第六百二十四条第一項の規定を準用する。

問い: 請負人は、仕事の目的物の引渡しを要する場合には、これを引き渡した後でなければ、報酬を請求することができない。

前者の文の特殊条件である「ただし」以下の文章を除けば、この2文の文末は「ない」が用いられている。しかし、前者の文は「義務」を表した文であり、一方で後者は「否定」を表している文であるため、この2文は偽の関係となる。つまり、「ない」に内在される意味を把握しないとこの違いを見抜くことは難しい。

このように問題を要素技術ごとに分類することで、複雑な問題構造を分解した検討ができ、今後の研究の新たな指針になる。

4. 自動解答タスクと人物関係抽出

平成29年度のテストデータは全69問あり、そのうち言語処理を行う際に、人物関係を抽出する必要のある問い及び人物役割を付与する必要のある問いは、それぞれ26問、27問存在した。しかし、現状の解答器では人物間の関係はほとんど考慮されておらず、表層的な処理だけとなっており、人物関係を抽出するタスクの解決は有用である。以下に人物関係を要する問いの例を挙げる。

甲土地を所有するAがBと通謀して甲土地をBに仮装譲渡し、AからBへの所有権移転登記がされた後、Bの債権者Cが甲土地を差し押さえた場合において、その差押えの時にCが仮装譲渡について善意であったときは、Aは、Cに対し、Bへの譲渡が無効であることを主張することができない。

この問いは正しい主張であるので、計算機は、Yesを出力する必要がある。この問いでは、A、B、Cという3人の人物が登場していることが明らかである。しかしこのアルファベットには人物のラベルという以上の意味がない。重要なのはこのA・Cの人物がどういう関係にあるのかということを経験が読み取ることである。ここでYesの根拠となる条文を示す。

第九十四条 相手方と通じてした虚偽の意思表示は、無効とする。

2 前項の規定による意思表示の無効は、善意の第三者に対抗することができない。

Aという人物をこの問いの主語とすると、Bは相手方、Cは善意の第三者という関係が読み取ることができる。よって、AがBに虚偽の意思表示である土地の仮装譲渡を行った後、(Aにとって)善意の第三者であるCに土地を差し押さえられたとしても、行為者Aは善意の第三者に対抗できないため、譲渡の無効を主張できない。

こうした人物関係について、動作主、被動作主、第三者などの属性をもつ辞書を構築することで、述語項構造解析による主格・目的格といったツールの出力と紐づけて自動的に人物関係を抽出することが可能になる。さらに、省略されている人物や事物の推測が容易になると考える。

5. 法律文書処理のための構造化辞書

先述の分類のうち、人物関係を対象として意味的關係を含めた構造化辞書を提案する。表3はH29年の司法試験のデータを基に作成した辞書の例である。一つの関係を一エントリとして表現する。関係そのものを表す行為とは、動作主と被動作主の立場を端的に説明する動詞である。また動作主・被動作主・第三者は、条文を基に抽出した。第三者とは、動作主と被動作主の関係を前提に、さらに何らかのアクションをとる存在である。これらの項目は、自然言語処理システムの利用を想定して実際の出現形を基本形にしたものを登録する。以下の条文を例に説明する。

第四百七十五条 弁済をした者が弁済として他人の物を引き渡したときは、その弁済をした者は、更に有効な弁済をしなれば、その物を取り戻すことができない。

第四百七十七条 前二条の場合において、債権者が弁済として受領した物を善意で消費し、又は譲り渡したときは、その弁済は、有効とする。この場合において、債権者が第三者から賠償の請求を受けたときは、弁済をした者に対して求償をすることを妨げない。

この場合は、動作主である「債権者」と被動作主である「弁済をした者」が関係そのものを表す行為「債権を有している」で結ばれている。「第三者」は、この関係を前提に、「第三者の行為」として、「債権者」に対して消費又は譲渡した物についての「賠償を請求」している。

今回は1年分のテストデータから抽出したが、エントリを増やすことで条文すべてを網羅できるであろう。さらには動作主・被動作主の属性を記述できると有用である。例えば、成年被後見人とは、精神上の障害により判断能力が欠けており、家庭裁判所から後見開始の審判を受けた人を指す。その成年後見人はこうした人々を後見する(法律行為などを代理したり補助したりする)人物である。こうした人物の可能な法律行為やそれぞれの関係性はエントリを増やすことによって得られるのではないかと。

6. 議論と展望

法律文書の辞書化における今後の課題をふたつ述べる。

抽象的な表現の扱いは難しい。例えば先の問いに対する条文では、「虚偽の意思表示」という抽象的な表現が用いられてい

た。具体的な記述がされた問いの文章から条文を検索するためには、問いの文章を抽象化する必要があるだろう。動詞に関しても同様である。「認める」「得る」「…がある」などといった一般性の高い動詞はよく用いられるが、解釈の幅が広く曖昧である。こうした動詞では動詞そのものよりもその項に具体的な表現がなされることが多いため、動詞そのものは利用されないことが多い。動詞の抽象化については、階層的に捉え、包含関係を表す手法[8]などが提案されているが、一般性が高いほど、動詞の置換は慎重な考慮が必要である。

もうひとつの課題は第三者の扱いについてである。先述のように、ここでの第三者には幅広く異なる立場のものが含まれる。権利を行使する立場であったり、許可をする立場であったりとその行為も多種多様である。次のステップとしては、行為の手がかりとして、表層格を用いた行為の判断を検討している。

7. おわりに

本研究では司法試験自動解答を念頭に、法律文書処理のリソース構築のため、司法試験の必要要素別分類を行った。さらにその中でも人物関係に着目して、人物関係の意味をとらえるための構造化辞書の設計と、手作業での辞書エントリサンプルの作成を行った。今後は作業やサンプルとする条文の数が増していく。これにより法律分野に現れうる単語を十分に網羅できる辞書構築を目標としている。こうしたリソースにより、自然言語処理による自動解答、自動処理に貢献できる。今後の研究では辞書を使った人物関係認識の自動化を実装する予定である。

謝辞 本研究は JSPS 科研費 17H06103, JSPS 科研費萌芽研究, JST CREST の助成を受けた。

参考文献

- [1] Chang, C. L. and Lee, R. C. T.. Symbolic Logic and Mechanical Theorem Proving. Academic Press, 1973, 331p.
- [2] M.-Y. Kim, R. Goebel, and S. Ken, "COLIEE-2015: Evaluation of Legal Question Answering," in Ninth International Workshop on Juris-informatics (JURISIN 2015), 2015.
- [3] M.-Y. Kim, R. Goebel, Y. Kano, and K. Satoh, "COLIEE-2016: Evaluation of the Competition on Legal Information Extraction/Entailment," in Tenth International Workshop on Juris-informatics (JURISIN 2016), 2016.
- [4] Y. Kano, M.-Y. Kim, R. Goebel, and K. Satoh, "Overview of COLIEE 2017," in the 16th International Conference on Artificial Intelligence (ICAAIL 2017), 2017.
- [5] 岡田拓哉, 竹内孔一, 石原靖弘: ニューラルネットワークを利用した日本語意味役割ラベル付与システムの構築, 言語処理, 2018
- [6] M.-Y. Kim and R. Goebel, "Two-step Cascaded Textual Entailment for Legal Bar Exam Question Answering," in the 16th International Conference on Artificial Intelligence and Law (ICAAIL 2017), 2017.
- [7] Y. Kano, R. Hoshino, and R. Taniguchi, "Analyzable Legal Yes/No Question Answering System using Linguistic Structures," in 4th Competition on Legal Information Extraction and Entailment (COLIEE 2017), 16th International Conference on Artificial Intelligence and Law (ICAAIL 2017), 2017.
- [8] 竹内孔一, 乾健太郎, 竹内奈央, 藤田篤: 意味の包含関係に基づく動詞項構造の細分類, <http://www.cl.cs.okayama-u.ac.jp/wp-content/uploads/2017/03/golex.pdf>, 2017

関係そのものを表す行為	その他とりうる行為	動作主	被動作主	第三者	第三者の行為
代理する, 代表する	責任を負う, 同意する	代理人, 法廷代理人	(本人), 未成年者, 制限行為能力者	家庭裁判所	認める
	同意を得る, 財産を処分する, 通知を発する	(本人), 未成年者, 制限行為能力者	代理人, 法定代理人		
保佐する・補助する	追認する,	保佐人・補助人 法定代理人	制限行為能力者	制限行為能力者の相手方	(制限行為能力者に)催告する, (保佐人・補助人に)行為を取り消す
所有する	償金を支払う, 償金を請求する, 通行する, 譲り渡す, 困障を設ける, 持分を取得する, 使用する, 分担する	主たる所有者 (本人)	従たる所有者, その他の所有者, その他の分割者		
債権を有する	取り立てる, (第三債務者に) 供託させる	質権者, 債権者	債務者	第三債務者	
	譲り渡す,	債務者		連帯債務者	更改する
	(動作主が) 取り戻す, (債権者が) 消費する, (債権者が) 譲渡する, (動作主に) 求償する	弁済をした者	債権者	第三者	物を引き渡される, 賠償の請求をする
売買する	代金の支払いを拒む, 契約の解除, 損害賠償の請求, 支払う	買主	売主		
	請求する, 契約の解除をする, 権利を移転する	売主	買主	他人	

表 3. 法律文書の人物関係構造化辞書のエントリ例