

権利義務認識のための契約書コーパスの構築

舟木 類佳^{1,2} 永田 祐介² 末永 幸平² 森 信介²¹ 株式会社 LegalForce ² 京都大学

ruka.funaki@legalforce.co.jp
 nagata.yusuke.88x@st.kyoto-u.ac.jp
 ksuenaga@fos.kuis.kyoto-u.ac.jp
 forest@i.kyoto-u.ac.jp

1 はじめに

契約書は複数の当事者が契約を交わす際に契約の内容、すなわち各当事者の権利や義務を定めるため用いられる。契約書の記述内容に不備があるとトラブルや訴訟問題に発展することもあるため、契約書を扱う法務担当者は契約時に念入りにレビューし、修正する必要があり、多大な時間がかかっているという課題がある。

そこで、契約書レビューのコスト軽減のために近年はコンピューターによるレビュー支援の取り組みが進んでいる。このような取り組みは様々なアプローチが存在するが、本研究では自然言語処理技術等を利用し契約書の理解を支援することを考える。

法務担当者が契約内容を確認する際には、自らに対して正しく権利が付与されているか、自らにとって不利な状況になっていないか、相手方に対して適切な義務が課されているか、などを正確に確認する必要がある。そこで、機械学習により権利や義務に関する情報を抽出し構造化したものを提示することができれば、契約書の理解を支援することが可能であると考えられる。それに先立ち、本研究では契約書の権利や義務の認識を目的としたコーパスを構築する。

契約書のコーパスを作成することは、弁護士などが持つ法務領域の知識が必要であるため、困難である。そこで、我々のコーパスは弁護士により作成された契約書を用い、弁護士のディレクションの下、契約書に対してアノテーションを施している。

本研究で扱うコーパスでは以下のような単語列の範囲を表すアノテーションを用いる。すなわち、(1) 契約書における当事者、(2) 各当事者が持つ権利、(3) 各当事者が持つ義務、(4) 権利や義務が効力を発動するための要件、(5) 権利や義務が効力を発動する例外、の5項目のアノテーションを行った。また、これらのアノテーション付与を行う2名の日英のアノテーターとともにアノテーター基準を定めた。

ラベル	説明
P	当事者 (Party)
R	権利 (Right)
O	義務 (Obligation)
C	要件 (Condition)
E	例外 (Exception)

表 1: アノテーションで用いるラベル

2 アノテーション

2.1 タグ

本研究におけるアノテーションでは、契約書に対して表 2.1 に表される 5 つのラベルを用いて XML のようなタグを付与する。タグを用いたアノテーションの文法は以下の通りである。

$$\begin{aligned}
 i, j, k &\in \mathbb{N} \\
 t &::= \langle tn \rangle \mid \langle /tn \rangle && \text{(tags)} \\
 tn &::= P_i && \text{(parties)} \\
 &\mid R_{j-p} && \text{(rights)} \\
 &\mid O_{k-p} && \text{(obligations)} \\
 &\mid C_{-rop} && \text{(conditions)} \\
 &\mid E_{-rop} && \text{(exceptions)} \\
 p &::= P_j \mid P_{j-p} \\
 rop &::= R_j \mid O_i \mid R_{j-rop} \mid O_{j-rop}
 \end{aligned}$$

タグは開始タグ $\langle tn \rangle$ に始まり、終了タグ $\langle /tn \rangle$ で閉じられる。ここで tn はタグ名を表す。また、タグで囲まれたテキストを内容と呼ぶ。入れ子構造であったり、範囲の重複は認めない。

タグ名については、開始タグと終了タグのペアで囲まれたテキストがどのようなラベルや対応関係を持つかを示す。すなわちタグ名は、それぞれ契約に関わる当事者、各当事者に与えられた権利、各当事者に課せられる義務、各権利が発動したり、各義務が課された

<P1>株式会社 ABC</P1> (以下「甲」という。) と <P2>株式会社 DEF</P2> (以下「乙」という。) は、甲が乙に開示又は提供する秘密情報の保持につき、次のとおり秘密保持契約 (以下「本契約」という。) を締結する。

図 1: 当事者のアノテーション例

りするための要件, 各権利や各義務に関する例外に対応する. 個別のタグ名の表し方について下記に詳細を述べる.

2.1.1 当事者

契約書は複数の関係者により締結される. 我々はこの契約に関わる関係者を当事者 (Party) と呼ぶ. 例えば「株式会社 ABC」と「株式会社 DEF」が秘密保持契約を結ぶ際には, それぞれ「株式会社 ABC」と「株式会社 DEF」が当事者である. 契約書中では当事者を指し示す略称が用いられることが多い (例えば甲, 乙, 売主, 買主, 等) が, 契約書のアノテーションにおいて, これらの用語は当事者としてアノテーションを行わず, はじめに登場する当事者名を一度だけアノテーションする.

当事者のタグはタグ名 P_i を持つ開始-終了タグ (自然数 i は ID と呼ぶ) でアノテーションする. ID i は異なる当事者を区別するために使用され, 契約書で登場した順番に番号が振られる. 例えば, 最初に登場した当事者には ID 1 が割り当てられ, 次の当事者には ID 2 が割り当てられる. ID は当事者を参照するために, 以降の契約書で利用される. 図 1 は当事者のアノテーションの実例である.

2.1.2 権利

契約書において各当事者の権利は「~することができる」などで表現される.

契約における当事者に権利が与えられている箇所は R_{j-p} というタグ名でアノテーションされる. ここで, p はハイフンで結合された当事者 P_i のリストである. 具体的には, p で表される当事者に対する権利は R_{j-p} というタグ名の開始タグと終了タグで囲まれた内容に対応する. ID j は異なる権利を区別するためにタグに付与され, これにより, この権利が行使されるための要件や例外のタグからの参照が可能となる. (2.1.4 節, 2.1.5 節参照). 図 2 は実際のアノテーションの例である.

2.1.3 義務

契約書において義務は「~しなければならない」, 「~するものとする」, 「~する」などで表される.

委託者は, <R6-P1>受託者に対して, 解約日の 1ヶ月前までに書面により通知することにより, いつでも本契約を解約する</R6-P1>ことができる.

図 2: 権利のアノテーションの例

乙は, <O7-P2>この契約により生じた権利・義務を第三者に譲渡してはならない</O7-P2>.

図 3: 義務のアノテーションの例

当事者 p の義務は $Ok-p$ というタグ名でアノテーションする. ID k は要件や例外のタグから参照する際に異なる義務として区別できるように付与されている. 図 3 は実際のアノテーションの例である. さらに図 4 は義務が複数の当事者に関連する例である.

2.1.4 要件

契約書の権利と義務が発動する要件は「~のとき」や「~の場合」などを用いて記述される.

要件のタグ名は $C-rop$ を用いる. ここで, rop は予め定義されている権利と義務への参照である R_j や Ok のハイフンで接続されたリストである. 図 5 は要件の実際の例を示す. また, 図 6 は一つの権利が複数要件をもつアノテーションの例である.

2.1.5 例外

契約書における権利と義務が発動する際の例外は「但し~の場合にはこの限りではない」や「~場合を除き」などの文言を使用して記述される.

例外のタグ名は $E-rop$ を用いる. ここで, rop は要件の場合と同様, 予め定義されている権利と義務への参照である R_j や Ok のハイフンで接続されたリストである. 図 7 は例外のアノテーションの実際の例である.

Remark 1 (コメント) 契約書におけるタイトルや条見出しなどは当事者における権利や義務のアノテーションと関連が無い. アノテーターがこれらの部分をコメントアウトできるよう文法を定義する. コメントの記号は#で示され, 行末までを無視することを表す.

2.2 アノテーションのガイドライン

アノテーションの一貫性が保たれるように, 次のようなガイドラインを定める.

1. 権利と義務の内容は主語の当事者を含まない.
2. 権利と義務のタグの内容はテキストの内容がわかる最小の単語列とする.

甲及び乙は、<O11-P1-P2>それぞれ本件業務に関する責任者を選任し、本契約締結後速やかに相手方に通知する</O11-P1-P2>ものとする。

図 4: 複数当事者に依存する義務のアノテーションの例

乙は、<C-O6>委託業務の遂行に際し甲に損害を与えた場合</C-O6>は、<O6-P2>速やかに損害賠償し</O6-P2>なければならない。

図 5: 要件のアノテーションの例

3. 権利と義務の内容は最大 1 つの動詞句を含む。もしも複数の動詞句同士が接続詞で接続されていれば、それぞれの句は分割される。
4. もしも否定を含む場合には、否定表現（「～しない」「～でない」等）もアノテーションに含む。
5. 内容は原則としては複数文に跨らない。但し、複数文に跨らなければ意味が通じないものに関しては例外的に複数文を跨いだアノテーションを認める。また、同一の主語をもつ同士が接続詞等で接続されている場合も認める。
6. 内容は文末にある句点（「。」）を含まない。

このガイドラインは日本語のアノテーションのものであり、英語のガイドラインも同様に作成した。英語の文法は日本語の文法といくつかの点で異なるものの、ガイドラインは類似の方法で作成可能である。

2.3 統計情報

英語契約書 46 件、日本語契約書 25 件に対して前節で示したアノテーションを行った。表 2 に統計情報を掲載する。

2.4 興味深い例

コーパスの作成に際し、いくつかの注目すべきアノテーションがあった。いくつかの例を紹介する。

図 8 は要件の対象が例外的に変化する例である。この例では列挙されている要件は全体的には譲渡人又は譲受人に関わるものだが、第 4 項は譲受人のみに例外的に限定されている。

図 9 は義務を負う当事者が契約締結時には定まらない例である。この例では、O10 の義務を負うべき当事者は、「解除をされた当事者」であり、P1, P2 のいずれであるかは、契約締結時には定まらない。本研究のアノテーションにおいては、このような場合には、

#1.
甲又は乙は、<C-R15>相手方に次の各号のいずれかに該当する事由が生じた場合</C-R15>には、<R15-P1-P2>何らの催告なしに直ちに本契約の全部又は一部を解除する</R15-P1-P2>ことができる。

- # (1)
<C-R15>重大な過失又は背信行為があった場合</C-R15>
- # (2)
<C-R15>支払いの停止があった場合、又は仮差押、差押、競売、破産手続開始、民事再生手続開始、会社更生手続開始、特別清算開始の申立があった場合</C-R15>
- # (3)
<C-R15>手形交換所の取引停止処分を受けた場合</C-R15>
- # (4)
<C-R15>公租公課の滞納処分を受けた場合</C-R15>
- # (5)
<C-R15>その他前各号に準ずるような本契約を継続し難い重大な事由が発生した場合</C-R15>

図 6: 一つの権利が複数の要件を持つアノテーションの例

いずれの本契約当事者も、<E-O10>本契約締結の事実、本契約の内容、本契約の締結に関連した交渉の経緯及び内容、本契約の締結に関連して相手方当事者から開示された秘密情報について、相手方当事者の事前の同意がある場合</E-O10>を除き、<O10-P1-P2>これを第三者に対して開示し、又は漏洩してはならず、また本取引以外の目的で使用してはならない</O10-P1-P2>。

図 7: 例外のアノテーション例

義務を負う可能性のあるすべての当事者を列挙する。本研究が目的とする権利義務認識においては、義務を負う可能性のある当事者をすべて挙げるのが望ましいため、このようなアノテーション基準とした。

図 10 は義務と要件が入れ子構造になる例である。このような場合は要件のアノテーションは行わないという選択をした。

3 関連研究

いくつかの権利や義務に関する認識の試みが存在する [2, 3, 1]。これらは我々の研究と権利と義務を認識するという点で関連しているが、いくつかの点で異なっている。まず、我々の研究がコーパス構築を目的としたアノテーション基準を議論するという点である。また、彼らのとっている方法が、文レベルの分類をベースにしており、我々が単語列からなる範囲を対象としているということである。また、どの当事者と関連するかということまで対応付けを行っている点も異なる。

Language	#契約書	#文	#P	#R	#O	#C	#E
English	46	6754	98	311	1050	498	72
Japanese	25	1994	49	92	438	268	7

表 2: コーパスの統計情報

譲渡人又は譲受人は、<C-R4>本事業譲渡の実行前に限り、以下の各号のいずれかが生じた場合</C-R4>には、<R4-P1-P2>相手方当事者に対し書面で通知することにより、本契約を直ちに解除する</R4-P1-P2>ことができる。ただし、<C-R5>第 4 号の場合</C-R5>については譲受人のみ<R5-P2>解除権を有する</R5-P2>ものとする。

#(1)

<C-R4>相手方当事者について本契約上の義務の不履行があり、相当期間を定めて履行を催告してもこれに応じない場合</C-R4>

#(2)

<C-R4>相手方当事者の表明及び保証の重要な部分につき、真実又は正確でないことが判明した場合</C-R4>

#(3)

<C-R4>相手方当事者について倒産手続等の開始又はその申立てがあった場合</C-R4>

#(4)

<C-R5>本契約締結後、譲渡資産、承継負債、本事業の経営状態、収益計画その他本事業に関し重大な悪影響を及ぼし得る事由があることが判明した場合</C-R5>

#(5)

<C-R4>上記各号のほか、本契約を継続し難い重大な事由が生じた場合</C-R4>

図 8: 例外の対象の当事者が例外的に変化する例

<C-O10>前項により解除が行われたとき</C-O10>は、解除をされた当事者は、<O10-P1-P2>相手方に対し負担する一切の金銭債務につき当然に期限の利益を喪失し、直ちに弁済し</O10-P1-P2>なければならない。

図 9: 当事者が定まらない例

4 おわりに

本研究では契約書の権利と義務の認識を目的としたアノテーション付きコーパスの構築を説明した。このコーパスは契約書の次のような範囲を示すアノテーションが施されている。(1) 契約に関わる当事者、(2) 当事者に与えられる権利、(3) 当事者に与えられる義務、(4) 権利や義務が効力を発揮する要件、(5) 権利や義務が効力を発揮する要件の例外、我々はアノテーション基準を定め、2名のアノテーターにより日英の契約書によりアノテーションを行った。

このコーパスは権利、義務の認識に用いることが可

本権利義務のうち、<O8-P1-P2>その移転又は対抗要件具備のために必要な登記、登録、通知、承諾その他の手続に要する登記費用その他一切の各当事者において発生する費用は、甲及び乙が別段の合意する場合を除き、各自の負担とする</O8-P1-P2>。

図 10: 義務の中で要件が記述されており、入れ子をさける為アノテーションを省略した例

能である。現在我々はこのコーパスを用いて機械学習により各当事者の権利、義務の認識に取り組んでおり、この問題を解くための適切な機械学習手法に関しても今後検討していきたい。

参考文献

- [1] I. Chalkidis, I. Androutopoulos, and A. Michos. Obligation and prohibition extraction using hierarchical RNNs. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 254–259, 2018.
- [2] I. Glaser, E. Scepankova, and F. Matthes. Classifying semantic types of legal sentences: Portability of machine learning models. *Proceedings of the International Conference on Legal Knowledge and Information Systems (JURIX)*, pages 61–70, 2018.
- [3] J. O’Neill, C. Robin, P. Buitelaar, and L. O’Brien. Classifying sentential modality in legal language: A use case in financial regulations, acts and directives. In *Proceedings of the International Conference on Artificial Intelligence and Law*, pages 159–168, 2017.