

# 日独民法における自動対応付け手法の比較と fine-tuning の実装

高橋寿記<sup>1</sup> 中村誠<sup>1</sup><sup>1</sup>新潟工科大学 工学部

202111096@cc.niit.ac.jp mnakamur@niit.ac.jp

## 概要

日本法と外国法において、類似条文の対応付けに一定の需要がある。先行研究 [1] では日本とドイツの民法の自動対応付けが行われた。しかし、対応付けできない条文も多く、高い評価は得られなかった。本研究では先行研究 [1] の対応付け手法と評価手法を見直し、fine-tuning の実装を行った。実験により、本研究で行った類似条文の自動対応付け手法が先行研究 [1] の手法を上回った。また、fine-tuning を実装することでさらに高い評価を得ることを確認した。

## 1 はじめに

日本と外国におけるビジネスにおいて、両国の法律について理解するのは重要なことである。特に、日本の法律では違法とされない場合でも、外国の法律では違法となる場合もあり、両国の法律の差異を知ることは必須である。

しかし、法律文書は一般的な文書に比べ、難しい表現や文体を用いることが多く、専門知識の無い人がすぐに理解することは難しい。さらに、外国の法律となると、言語や法体系等の知識も必要になる。図 1 に実際の日本とドイツの民法の条文、ドイツ民法の翻訳文書をそれぞれ示す。図 1 は実際に対応関係にある条文であるが、ドイツ語の知識が無いと何が書かれているか分からない上に、日本の条文であっても専門的な用語があり理解が難しい。しかし、実際の法律を読んで両国の法律の差異を理解するには、このような言語の違う数百の条文の中から対応する条文を見つけなければならない。そのためには、外国の法律の全ての条文を読み、どの条文にどのようなことが書いてあるか理解する必要がある。

本研究では日本の法律と外国の法律を読み込み、日本のある条文が外国のどの条文に類似しているか自動で判定し、その類似条文の対応付けを行うシステムを開発している。これにより、専門知識が無い

人でもシステムの出力を見れば、一目でどの条文が対応しているか理解できる。

先行研究 [1] では、日本とドイツの民法における類似条文の自動対応付け手法として、条文単位による 1 対 1 の対応付けが行われた。しかし、実際の対応関係では、1 対多や多対多の関係が存在する。そのため、対応付けできない条文も多くあり、モデルの評価は高くなかった。

民法（明治 29 年法律第 89 号）

第 726 条 親等は、親族間の世代数を数えて、これを定める。

2 傍系親族の親等を定めるには、その一人又はその配偶者から同一の祖先にさかのぼり、その祖先から他の一人に下るまでの世代数による。

### (a) 日本民法の原文

Bürgerliches Gesetzbuch (BGB)

§ 1589 Verwandtschaft

(1) Personen, deren eine von der anderen abstammt, sind in gerader Linie verwandt. Personen, die nicht in gerader Linie verwandt sind, aber von derselben dritten Person abstammen, sind in der Seitenlinie verwandt. Der Grad der Verwandtschaft bestimmt sich nach der Zahl der sie vermittelnden Geburten.

### (b) ドイツ民法の原文

ドイツ民法典第 4 編（親族法）

第 1589 条 血族関係

(1) 一方の者が他の者と出自関係にあるときは、これらの者は、直系の血族関係にある。直系の血族ではないが、同一の第三者と出自関係にあるときは、これらの者は、傍系の血族関係にあるものとする。血族の親等は、血族関係を基礎づける出生の数により定まる。

### (c) ドイツ民法の翻訳文書

## 図 1 日本民法とドイツ民法の比較

また、法律文書は似た表現が多く、対応条文と非対応条文を既存のモデルにより判別するのは難しい。

したがって、本研究の目的は、条文単位による 1 対多や多対多の自動対応付けを実装するとともに、類似条文の対応付けに特化した fine-tuning モデルの有効性を検証することである。

以降、第 2 節では本研究の背景、第 3 節では実験、第 4 節では実験結果、第 5 節では考察、第 6 節は本研究についてのまとめを述べる。

## 2 提案手法

図 2 に先行研究 [1] の対応付け手法と、本研究の対応付け手法の例を示す。

先行研究 [1] では、類似条文の自動対応付け手法として、日本とドイツの条文間において、双方向で最も類似度が高くなった組を対応付けた。図 2 の赤の矢印は、日本民法第 726 条とドイツ民法第 1589 条が対応付けられた様子を表している。

本実験ではデータセットで与えられた日本とドイツの条文の全てのペアで類似度を求め、定めた閾値以上の類似度だったペアを対応付けることとした。

図 2 より、閾値  $\theta$  が 0.90 であったとき、日本民法の第 726 条の条文に対して、ドイツ民法の全ての条文との類似度を計算すると、第 1589 条に加えて、第 1590 条が類似度 0.90 以上であるため対応付けられる。青の矢印は、本実験の手法で新たに日本民法第 726 条とドイツ民法第 1590 条が対応付けられた様子

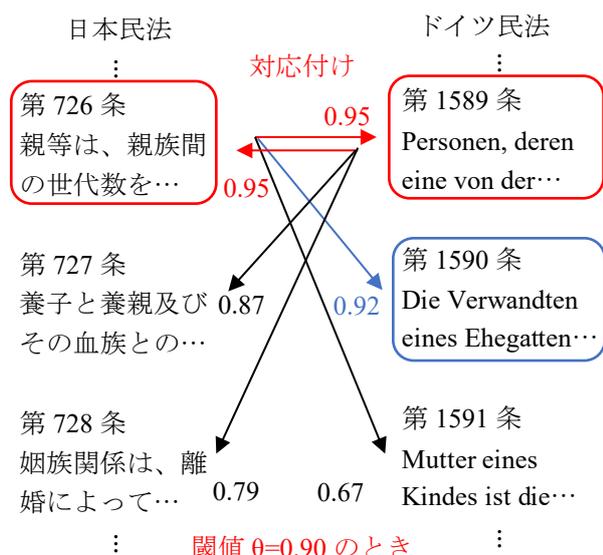


図 2 対応付け手法

を表している。

これにより、先行研究 [1] で問題であった日本の条文 1 つに対し、正解条文が複数ある場合でもドイツ条文 1 つしか対応付けできないという問題を解消できる。

類似度の算出方法は先行研究 [1] と同じく、ベクトル化した条文のコサイン類似度を求める。

対応付けでは、日本民法とドイツ民法の翻訳文書を汎用日本語事前学習モデルにより対応付ける方法（以降、日本語対応付けという。）と、日本とドイツの民法を原文のまま多言語モデルで対応付ける方法（以降、ドイツ語対応付けという。）の 2 通りを行った。

## 3 実験

### 3.1 法律データ

日本民法は、デジタル庁が管轄する法令検索サイト e-Gov<sup>i</sup>より民法（明治 29 年法律第 89 号）を xml 形式で入手し、利用した。

ドイツ民法は、ドイツの法令を扱う連邦司法・消費者保護省の公式 web ページ<sup>ii</sup>から Bürgerliches Gesetzbuch (BGB)を xml 形式で入手し、利用した。

ドイツ民法の翻訳文書は、法務省大臣官房司法法制部が作成した翻訳文書<sup>iii</sup>を pdf 形式で入手し、利用することとした。

図 1 (a) の日本民法と図 1 (c) の翻訳文書を比較すると、翻訳文書は法律の専門家が執筆していることから、日本民法と文体が似ていることが分かる。

日本語対応付けでは、データセットとして、日本民法より第 725 条から第 881 条と、ドイツ民法の翻訳文書より第 1297 条から第 1888 条について、全ての組み合わせで作成した 86,925 個のデータセットを使用した。そのうち、学習データに 72,438 個、テストデータとして 14,487 個を使用した。

ドイツ語対応付けでは、データセットとして、日本民法より第 725 条から第 881 条と、ドイツ民法の原文より第 1297 条から第 1888 条について、全ての組み合わせで作成した 86,010 個のデータセットを使用した。そのうち、学習データに 71,675 個、テストデータとして 14,335 個を使用した。

正解データは先行研究 [1] と同じく、民法に関す

<sup>i</sup> <https://www.e-gov.go.jp/>

<sup>ii</sup> <https://www.gesetze-im-internet.de/index.html>

<sup>iii</sup> <https://www.moj.go.jp/content/001387121.pdf>

る注釈本 [2-4] を元に、人出で作成されたものを用いた。正解データはデータセットに合わせ、日本民法の第 725 条から第 881 条までの部分を使用した。

### 3.2 評価手法

先行研究 [1] では、モデルの評価として、対応付け結果から TP, FP, TN, FN を集計し、Precision, Recall, F1 値で評価を行った。その集計方法を表 1 に示す。日本の条文に対し、正解のドイツ条文を対応付けた際は TP, 対応関係のない条文を対応付けなかったとき TN, 間違った条文を対応付けたとき FP, 対応関係のある条文を対応付けなかったとき FN として集計を行った。

表 1 先行研究[1]の集計方法

日本	システム出力 (ドイツ)	正解 (ドイツ)	判定
⋮	⋮	⋮	⋮
726	1589	1589, 1590	TP
727			TN
728	1600	1590	FP
729		1764	FN
⋮	⋮	⋮	⋮

本実験では、評価指標は先行研究 [1] と同じものを用いた。しかし、対応付け手法の変更に伴い、集計方法の改善を行った。その集計方法を表 2 に示す。

表 2 の日本の第 726 条に対してドイツの第 1589 条, 第 1590 条, 第 1591 条, 第 1592 条のように、表 1 と違い、日本の条文 1 つに対し、与えられたペア全てにおいてシステムは判定を行っている。ここで、出力が対応で正解と一致すれば TP, 対応で不一致な

表 2 本実験の集計方法

日本	ドイツ	出力	正解	判定
⋮	⋮	⋮	⋮	⋮
726	1589	対応	対応	TP
726	1590	対応	対応	TP
726	1591	非対応	非対応	TN
726	1592	対応	非対応	FP
⋮	⋮	⋮	⋮	⋮
727	1589	非対応	非対応	TN
⋮	⋮	⋮	⋮	⋮

ら FP, 非対応で一致なら TN, 非対応で不一致なら FN となる。

### 3.3 問題設定

日本語対応付けでは、BERT [5] の日本語版である汎用日本語事前学習モデル<sup>iv</sup> (以降, 東北大 BERT という。)を用いた。また、ドイツ語対応付けでは、多言語モデルとして MiniLM<sup>v</sup>を使用した [6]。

本実験では fine-tuning に 5 分割交差検定を用いた。日本語対応付け, ドイツ語対応付けともに、学習データを 5 分割し、4 個を訓練データ, 残りを検証データとして学習した。

日本語対応付け, ドイツ語対応付けともに、fine-tuning の際にはデータセットにラベルが付与されている。このラベルは、正解データにおいて対応関係にある条文ペアは「1」とし、それ以外のペアは「ペアの類似度×定減率」とした。条文ペアの類似度は、日本語対応付けでは東北大 BERT を使い、ドイツ語対応付けには MiniLM を用いて条文をベクトル化し、コサイン類似度を求めた。

### 3.4 実験方法

本研究では以下のような実験を行った。各実験は、日本語対応付けとドイツ語対応付けの両方でを行い、それぞれ評価を行う。

**[実験 1]** fine-tuning に適した定減率の調査

**[実験 2]** 先行研究 [1] の対応付け手法で対応付けを行い、本研究の評価手法を用いる場合

**[実験 3]** fine-tuning を行わずに提案手法で対応付けし、本研究の評価手法を用いた場合

**[実験 4]** fine-tuning を行い提案手法で対応付けし、本研究の評価手法を用いた場合

fine-tuning を行うにあたり、データセットのうち、対応関係の無い条文ペアのラベルとして「ペアの類似度×定減率」を用いる。このラベルについては、モデルが非対応条文の類似度を低く学習するように、適切な定減率  $r$  を定める必要がある。

実験 1 では定減率  $r$  を  $0.05 \leq r \leq 1.0$  の範囲で 0.05 ずつ変化させて繰り返し fine-tuning を行い、最も F1 値が高くなる定減率  $r$  を調査することを目的とする。

実験 2 では、先行研究 [1] の 1 対 1 の対応付け手法を用いたとき、1 対多の評価手法では評価がどの

<sup>iv</sup> cl-tohoku/bert-base-japanese-whole-word-masking

<sup>v</sup> sentence-transformers/paraphrase-multilingual-MiniLM-L12-v2

ように変化するか調査することを目的とする。

実験 3 では、1 対多の対応付けが可能な本研究の手法が、先行研究 [1] の 1 対多の対応付けが不可能な手法に比べて、評価が高くなるか調査することを目的とする。

実験 4 では、1 対多の対応付けに加え、対応付けタスクに特化した fine-tuning モデルにより、評価が向上するか調査することを目的とする。

本研究の対応付け手法では、閾値  $\theta$  を元に対応、非対応を判定するため、適切な値を定める必要がある。本実験では閾値を  $0 < \theta \leq 1$  の範囲で 0.01 ずつ変化させて対応付けを行い、F1 値が最大となったときの値を最適な閾値であるとし、そのときの評価を結果として記録した。

## 4 実験結果

実験 1 の結果を図 3 に示す。ここで、日本語対応付けを青のグラフ、ドイツ語対応付けをオレンジのグラフとしている。

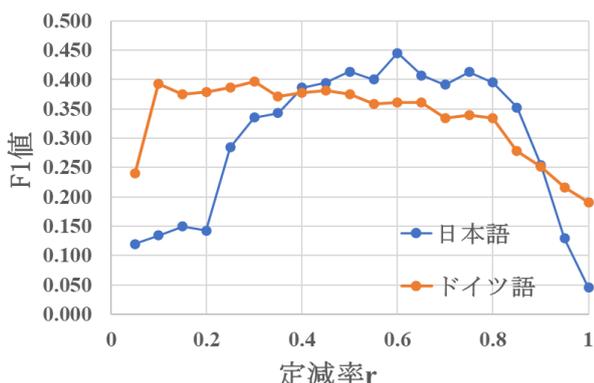


図 3 定減率  $r$  の変化による F1 値の変化

図 3 より、日本語対応付けでは定減率 0.60 のときに 0.4457 で最大となっている。ドイツ語対応付けでは定減率 0.30 のときに 0.3965 で最大となっている。

次に、実験 2 の結果を表 3 に示す。ここで、日本語は日本語対応付け、ドイツ語はドイツ語対応付けを表す。また、先行研究 [1] の対応付け手法と評価手法では、日本語対応付けで F1 値が 0.0172、ドイツ語対応付けで F1 値が 0.2238 であった。

実験 3 と実験 4 の結果は、表 4 に示す。ここで、日本語 ft とドイツ語 ft は fine-tuning を行った際の結

表 3 先行研究 [1] の手法による対応付け (実験 2)

	Precision	Recall	F1 値
日本語	0.0455	0.0036	0.0067
ドイツ語	<b>0.2807</b>	<b>0.0597</b>	<b>0.0985</b>

果を示し、日本語 ft は定減率 0.60、ドイツ語 ft は定減率 0.30 の F1 値が最も高くなったときの値を示す。また、fine-tuning の評価結果は 5 分割交差検定で作成した 5 つのモデルの平均となっている。

表 4 本研究の手法による対応付け (実験 3, 4)

	Precision	Recall	F1 値	$\theta$
日本語	0.0172	0.0569	0.0264	0.96
日本語 ft	0.6269	<b>0.3713</b>	<b>0.4457</b>	0.66
ドイツ語	0.1044	0.1289	0.1152	0.76
ドイツ語 ft	<b>0.6401</b>	0.2894	0.3965	0.37

## 5 考察

表 3 と表 4 を比べると、本研究の評価手法においては、先行研究 [1] よりも本研究の対応付け手法の方が高い評価が得られた。これは 1 対多の対応付けが可能となり、1 対 1 の対応付けの時よりも多くの条文を対応付けることが出来たためだと考えられる。また、fine-tuning 実装前に比べて、実装後の方が大幅に評価は高くなっていることが分かる。

日本語対応付けとドイツ語対応付けを比較すると、表 3 の先行研究 [1]、表 4 の本研究の fine-tuning なしの対応付け手法では、ドイツ語対応付けの方が評価は高い。これは、東北大 BERT の日本語空間によるベクトル化では、条文の文体が似ており、対応と非対応に関係なく条文間の類似度が高くなったため、評価が低くなったと考えられる。

また、表 4 の fine-tuning 後の結果で日本語対応付けの方が高くなったのは、元々多言語空間のベクトル化で類似度に差が出ていた MiniLM よりもほとんど差が無い東北大 BERT の方が、対応ペアの類似度を高く、非対応ペアの類似度を低く学習する fine-tuning の効果が高かったためだと考えられる。

## 6 おわりに

本研究では、日本の条文 1 つに対し、ドイツの条文を複数対応付けできる手法を提案した。また、fine-tuning により対応付け性能の向上を試みた。結果として、日本とドイツの条文を一体一で対応付ける先行研究 [1] に比べ、本研究の手法の方が評価は高くなった。また、fine-tuning の実装により評価が大幅に向上した。今後は、対応付けの評価をさらに向上させるため、fine-tuning に使用するデータセットの見直しなどを行っていきたい。

## 謝辞

本研究は JSPS 科研費 JP24K03230 の助成を受けたものです。

## 参考文献

- [1] 山田大地, 中村誠 (2024). 法律間の類似条文の対応付けにおける BERT の法令ドメイン適応. 言語処理学会第 30 回年次大会発表論文集, pp. 1334-1338.
- [2] 二宮周平(編) (2017). 新注釈民法(17)親族(1) 725 条～791 条. 有斐閣.
- [3] 中川善之助, 山畠正男(編) (1994). 新版注釈民法(24) 親族(4) 792 条～817 条の 11. 有斐閣.
- [4] 於保不二雄, 中川淳(編) (1994). 新版注釈民法(25)親族(5) 818 条～881 条. 有斐閣.
- [5] Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova (2019). “BERT: Pre-training of deep bidirectional transformers for language understanding.” In Proceedings of NAACL, pp. 4171-4186.
- [6] Wenhui Wang, Furu Wei, Li Dong, Hangbo Bao, Nan Yang, Ming Zhou (2020). “MiniLM: Deep Self-Attention Distillation for Task-Agnostic Compression of Pre-Trained Transformers.” arXiv:2002.10957