T2R2 東京科学大学 リサーチリポジトリ
Science Tokyo Research Repository

論文 / 著書情報
Article / Book Information

論題	判決書要約文の自動評価
著者	新保彰人, 山田寛章, 徳永健伸
出典	言語処理学会第31回年次大会(NLP2025)発表論文集, pp. 1974-1979
発行日	2025, 3
権利情報	Creative Commons Attribution 4.0 International License(https://creativecommons.org/licenses/by/4.0/).

Powered by T2R2 (Science Tokyo Research Repository)

判決書要約文の自動評価

新保彰人　山田寛章　徳永健伸
東京科学大学　情報理工学院
shimbo.a.aa@m.titech.ac.jp, yamada@comp.isct.ac.jp, take@c.titech.ac.jp

概要

一般の文書要約で使われている ROUGE などの自動評価指標は判決書自動要約タスクでも使われている．しかし既存の自動評価指標では判決書要約文に不可欠な要素が，要約文に含まれているかを評価することができない．本研究では，判決書要約文に特化した評価ルーブリックを策定し，それに基づいて法律の専門家による人手評価を行う．そして，その評価データを利用して判決書要約文に特化した自動評価器を構築する．構築した評価器をカッパ係数で評価し，自動評価器と正解データとの一致度が人手評価者の間の一致度を部分的に上回ることを示す．

1　背景

日本において，民事判決のオープンデータ化の構想が進んでいる．現状では裁判所が公開する民事判決の数は限られているが，オープンデータ化によってその数は増大すると考えられる [1]．

大量の判決書を効率的に活用する方法の一つとして自動要約が挙げられる．通常，判例集などに掲載される判決書には，判示事項と呼ばれる判決書の要約文に相当するテキストが付されている．判示事項は法律の専門家が判決書を検索したり，判決書の内容を素早く把握する際に利用される．現状では判示事項は人手で記述されているが，大規模言語モデルで判示事項を自動生成することで，大量の判決書の効率的な利用が可能になると期待されている．

法律系文書の自動要約の研究は既に存在するが，既存の研究では生成要約文の評価に ROUGE [2] や BERTScore [3] などの自動評価指標を利用している [4, 5]．判示事項には，裁判における原告と被告の主張が対立している部分である「争点」が含まれることが望ましいとされるが，既存の自動評価指標では「争点を含むか」のような判示事項特有の観点での評価ができないという課題がある．

本研究では，法律の専門家と協力して判決書要約文に特化した評価ルーブリックを策定する．そのルーブリックに基づく人手評価を行い，その評価データを利用して LLM ベースの自動評価器を構築する．さらに自動評価器と人手評価の一致度と，異なる評価者同士の人手評価の一致度の比較を行う．

2　関連研究

Calzolari ら [6] は刑事訴訟判例の要約文評価の手法を提案した．刑事訴訟において，犯行の意図は罪の成立要件に関わる重要な要素である．この点に着目し Calzolari らは要約文に犯行の意図を示す表現が過不足なく含まれているか，という情報に基づく評価手法を提案した．本研究で扱う民事訴訟判例においても，争点のような民事判例において重要な要素が要約文に含まれるかどうかが要約の質を左右する．本研究で策定するルーブリックでは 8 種類の評価観点を設定するが，その内の 5 種類はある要素を含むかどうかに関する観点である．

近年では LLM を利用した要約評価の研究も行われている．Siledar ら [7] は意見要約文の 7 種類の評価観点に基づく評価値を出力する自動評価器を LLM を利用して構築した．Siledar らは各評価観点に対し，5 段階で評価基準を定めた．Siledar らの手法はプロンプトに評価基準の内容を入力する手法であり，既存の評価指標よりも人手との相関が高いことが示されている．本研究においてもルーブリックの内容をプロンプトに入力する手法を実験する．

3　データセット

3.1　生成要約文

判決書要約文の人手評価を行うために，評価対象となる要約文を自動要約器で生成する．自動要約器は Swallow-MX-8x7b-NVE-v0.1 [8] を QLoRA [9] でファインチューニングして構築する．ファインチューニングには株式会社 LIC によって提供された民事通常訴訟事件の判決書と判示事項の組 4,899 件

This work is licensed by the author(s) under CC BY 4.0
(https://creativecommons.org/licenses/by/4.0/).

を利用する．構築した自動要約器で 400 件の判決書を要約し，その要約文を人手評価用のデータとして使う．表 6 に生成要約文と，人手で書かれた判示事項の例を示す．

3.2 評価ルーブリック

人手評価と自動評価の評価基準となるルーブリックを表 1 に示す．評価観点は 8 種類あり，各観点は 3 段階で評価する．このルーブリックは法律の専門家と協力して定めた．

ルーブリック中の**事案の概要**とは当事者（被告と原告）の関係や紛争に至る背景を説明したものである．**争点**とは，当事者の主張の間で相違が生じている点のことであり，**争点に対する判断**とは争点に対して裁判所が下した判断のことである．争点が複数ある場合にはそれぞれの争点に対する判断が下される．**請求**とは原告が被告に対して請求している行為のことである．例えば，賠償金の支払いの請求もこれに該当する．**請求に対する結論**は原告の請求に対して裁判所が下した判断のことであり，請求を認容したか，あるいは棄却したかが主な内容になる．判例によっては，請求の一部を認容し一部を棄却する場合もある．

3.3 人手評価

3.1 で生成した要約文 400 件に対してルーブリックに基づく人手評価を実施する．また，ルーブリックに基づく評価の他に，要約文に対するコメントを自由記述する．

評価者は合計で 8 名で，4 ペアに分かれて評価する．要約文 400 件を 100 件ずつに分け，それぞれを異なるペアに割り当てる．各ペアは最初に，各自が独立して評価を実施し，その後ペア内で合議をして最終的な評価を決定する．

収集した合議後の評価データ 400 件を学習セット 160 件，開発セット 120 件，テストセット 120 件に分割する．

4 手法

収集した人手評価データを用いて，判決書要約文評価器を構築する．評価器のベースモデルには OpenAI 社が提供する gpt-4o-mini-2024-07-18 [10] を利用する．学習と推論は OpenAI API を利用して行う．プロンプトを表 7 に示す．

手法は Zero-shot, Few-shot, Fine-tuning, Fine-tuning-with-comments の 4 つである．

Zero-shot では学習を行わずモデルに判決書と要約文のみを入力し，評価スコアを生成させる．

Few-shot では，評価対象の判決書と要約文の前にサンプルとして 5 件の判決書と要約文と人手評価の三つ組を入力する．サンプル 5 件を選定する際には，各評価観点について，スコア 0，1，2 のすべてがサンプル 5 件の人手評価スコアの中に 1 つ以上含まれるようにした．表 2 に Few-shot に使ったサンプルの人手評価の値を示す．

Fine-tuning では学習セットと開発セットを使って人手評価を学習する．

Fine-tuning-with-comments でも学習を行うがこの手法では判決書と生成要約文を入力し，評価コメントと評価スコアを生成する学習を行う．この手法は Chain-of-Thought（CoT） [11] から着想を得ている．CoT では LLM がタスクを解く際の思考の過程を生成させ，最終的なタスクの精度を向上させる．例えば，算数の文章問題を解くタスクで途中式を出力させることで精度が向上することが報告されている [11]．Fine-tuning-with-comments では評価スコアの根拠としてコメントを生成させることで通常の Fine-tuning よりも精度が向上することを期待する．

学習を行う手法ではエポック数を 3 に設定する．モデルの学習には学習セットと開発セットを使い，テストセットで手法を評価する．ただし手法の探索の過程では学習セットのみでモデルの学習を行い，開発セットでの評価を行なった．

モデルの評価は各評価観点ごとに Quadratic Weighted Kappa（QWK） [12] で行う．

5 結果

各手法の生成スコアの正解スコアに対する一致度を表 3 に示す．また，評価器のパフォーマンスの比較対象として各ペア内の合議前の評価スコアの一致度を QWK で評価した結果を表 4 に示す．ただし，表 3 の値はテストセット 120 件における生成スコアと**合議後**人手スコアを比較しているのに対し，表 4 の値は 400 件全ての**合議前**人手スコアを比較している点に留意されたい．

5.1 手法の比較

Zero-shot は「④請求に対する結論を含む」以外で，Few-shot は「③争点に対する判断を含む」以外で QWK が 0.2 を下回った．Landis ら [13] の基準（表 5

This work is licensed by the author(s) under CC BY 4.0
(https://creativecommons.org/licenses/by/4.0/).

表 1　判決書要約文評価ルーブリック

評価観点	スコア 2	スコア 1	スコア 0
① 事案の概要	含んでいる。	含んでいるが内容が不十分。	含んでいない。もしくは記述されているが判決書本文とは全く関係がない。
② 争点を含む	含んでいる。	含んでいるが内容が不十分。	含んでいない。もしくは記述されているが判決書本文とは全く関係がない。
③ 争点に対する判断を含む	含んでいる。	含んでいるが内容が不十分。	含んでいない。もしくは記述されているが判決書本文とは全く関係がない。
④ 請求に対する結論を含む	含んでいる。	含んでいるが内容が不十分。	含んでいない。もしくは記述されているが判決書本文とは全く関係がない。
⑤ 判決書に対して矛盾を含まない	矛盾がない。	矛盾ではないが、判決書に記述されていない内容を含む。	矛盾を含む。
⑥ 文法上適切である	問題なし。	判示事項として不自然な表現を含む。（です・ます調など）	文法的に正しくない表現を含む。もしくは文として破綻している。
⑦ 適切な長さである	問題なし。	実用に耐えうるが、より簡潔に記述するべきである。	実用に耐えない程度に長い。
⑧ 論理的な一貫性がある	論の展開や記述の順番に破綻がない。	実用に耐えうるが、論理の展開や記述の順番が不自然である。	論理の展開や記述の順番に破綻がある。

表 2　Few-shot 用サンプルの人手評価スコア

評価基準	① 事案の概要を含む	② 争点を含む	③ 争点に対する判断を含む	④ 請求に対する結論を含む	⑤ 判決書に対して矛盾を含まない	⑥ 文法上適切である	⑦ 適切な長さである	⑧ 論理的な一貫性がある
サンプル 1	2	2	2	2	2	2	2	2
サンプル 2	2	0	0	0	0	0	0	0
サンプル 3	1	1	1	1	1	1	1	1
サンプル 4	0	2	1	2	1	2	2	1
サンプル 5	1	1	1	2	2	2	1	2

参照）では 0.0〜0.2 は「わずかな一致」とされる．

Few-shot は 5 観点で Zero-shot を上回っているものの，残りの 3 観点では下回り，本タスクにおいて Few-shot の有用性は確認できなかった．

「①事案の概要を含む」と「⑥文法上適切である」は Fine-tuning が最良となり，それ以外の 6 観点は Fine-tuning-with-comments が最良となった．大半の評価観点において，モデルにスコアの根拠となるコメントを生成させることでスコア予測の精度が高まることが確認された．

5.2　評価器と人手評価の比較

表 3 の Zero-shot 及び Few-shot と表 4 の各ペアの一致度を比較する．Zero-shot は「③争点に対する判断を含む」と「④請求に対する結論を含む」においてペア A を上回っているがそれ以外では全てのペアの一致度を下回っている．Few-shot にも同様の傾向が見られる．このため，Zero-shot と Few-shot の性能は人間に劣ると考えられる．

最も精度が高い Fine-tuning-with-comments と人手の各ペアの一致度を比較する．「④請求に対する結論を含む」は 4 ペア全ての一致度を評価器が上回った一方で，「⑥文法上適切である」は 4 ペア全ての一致度を下回った．それ以外の観点では評価器が上回るペアと下回るペアが混在する結果となった．この結果から Fine-tuning-with-comments は人間とほぼ同程度の性能であると考えられる．

5.3　人手評価の分析

表 4 中に示されたカッパ係数の値の分布を表 5 に示す．「まずまずの一致」から「中程度の一致」に該当するものが多いが，「わずかな一致」に該当する

This work is licensed by the author(s) under CC BY 4.0
(https://creativecommons.org/licenses/by/4.0/).

表 3　正解スコアに対する一致度（太字は最高値）

手法 \ 評価基準	①事案の概要を含む	②争点を含む	③争点に対する判断を含む	④請求に対する結論を含む	⑤判決書に対して矛盾を含まない	⑥文法上適切である	⑦適切な長さである	⑧論理的な一貫性がある
Zero-shot	0.082	0.146	0.198	0.210	0.022	0.020	-0.032	0.008
Few-shot	0.086	0.196	0.203	0.019	0.000	0.009	0.072	0.013
Fine-tuning	**0.403**	0.136	0.211	0.573	0.101	**0.296**	0.332	0.168
Fine-tuning-with-comments	0.384	**0.387**	**0.263**	**0.591**	**0.243**	0.215	**0.361**	**0.320**

表 4　合議前の人手評価の一致度

ペア \ 評価基準	①事案の概要を含む	②争点を含む	③争点に対する判断を含む	④請求に対する結論を含む	⑤判決書に対して矛盾を含まない	⑥文法上適切である	⑦適切な長さである	⑧論理的な一貫性がある
ペア A	0.402	0.216	0.145	0.196	0.327	0.235	0.085	0.222
ペア B	0.443	0.463	0.407	0.299	0.262	0.328	0.564	0.277
ペア C	0.147	0.150	0.285	0.311	0.184	0.220	0.397	0.183
ペア D	0.423	0.443	0.370	0.411	0.362	0.420	0.429	0.441

ものも 7 件見られた．また，「かなりの一致」以上に該当する 0.6 以上のカッパ係数は見られなかった．この結果から，本研究で行なった評価作業は専門家の間でも意見が分かれるタスクであったといえる．特に「⑤判決書に対して矛盾を含まない」は全てのペアで 0.4 を下回っており，矛盾に関する評価は意見の相違が生じやすいものであったとみられる．

今回使用したルーブリックは比較的簡潔に記述されており，評価者間でルーブリックの解釈に差異が生じていた可能性がある．これにより評価者間の評価スコアの不一致が生じたと考えられる．

信頼性の高い人手評価データを収集するためには，ルーブリック策定の段階で評価基準を詳細に定義することが有効であると考えられる．

表 5　表 4 のカッパ係数の分布

範囲	Landis ら [13] の基準	度数
$0.0 < \kappa \leq 0.2$	わずかな一致	7
$0.2 < \kappa \leq 0.4$	まずまずの一致	14
$0.4 < \kappa \leq 0.6$	中程度の一致	11
$0.6 < \kappa \leq 0.8$	かなりの一致	0
$0.8 < \kappa \leq 1.0$	ほとんど完全一致	0

6　結論

本研究では判決書要約文の人手評価と，自動評価器の構築を行なった．Zero-shot や Few-shot では高い精度が確認できなかったが，評価器にスコアとともにコメントを生成させる学習をすることにより精度が向上することが確認された．評価器の生成スコアと正解スコアの一致度は部分的に人手評価同士の一致度を上回り，特に「④請求に対する結論を含む」という観点では全てのペアの一致度を上回った．

今後の展望として，評価器のさらなる精度向上と要約器の精度向上への応用が挙げられる．

評価器の精度向上の方法として，評価ルーブリックの詳細化が挙げられる．本研究では表 1 のルーブリックを人手評価と自動評価器のプロンプトに共通して使用したが，人手評価の際には各評価者がルーブリックより詳細な評価基準を持って作業を行なっていた可能性がある．したがって，評価者の評価基準を言語化してプロンプトに記述することで評価器の精度が向上すると考えられる．また，詳細化したルーブリックで再度人手評価を実施すれば信頼性の高いデータを収集できると考えられる．そのようにして収集したデータで評価器を構築すればさらに高い精度の評価器になる可能性がある．

本研究で構築した評価器を要約器の評価に利用することで要約性能の改善が期待できる．例えば生成要約文を自動評価し，その評価が高いものを使って再び要約器の学習することで要約精度が向上すると考えられる．また，強化学習における報酬関数として自動評価器を利用することも可能である．

今後は以上のような評価器及び要約器改善の手法を探求する予定である．

This work is licensed by the author(s) under CC BY 4.0
(https://creativecommons.org/licenses/by/4.0/).

謝辞

本研究は株式会社 LIC の支援を受けたものです.

参考文献

[1] 公益財団法人日弁連法務研究財団. 民事判決情報のオープンデータ化に向けた取りまとめ. https://www.jlf.or.jp/wp-content/uploads/2021/04/pt-houkoku20210325.pdf. Accessed: 2024-12-21.

[2] Chin-Yew Lin. ROUGE: A package for automatic evaluation of summaries. In **Text Summarization Branches Out**, pp. 74–81, Barcelona, Spain, July 2004. Association for Computational Linguistics.

[3] Tianyi Zhang, Varsha Kishore, Felix Wu, Kilian Q. Weinberger, and Yoav Artzi. Bertscore: Evaluating text generation with BERT. In **8th International Conference on Learning Representations, ICLR 2020, Addis Ababa, Ethiopia, April 26-30, 2020**. OpenReview.net, 2020.

[4] 新保彰人, 菅原祐太, 山田寛章, 徳永健伸. 大規模言語モデルを用いた日本語判決書の自動要約. 言語処理学会第 30 回年次大会, Kobe, Japan, March 2024.

[5] Yang Zhong and Diane Litman. STRONG – structure controllable legal opinion summary generation. In Jong C. Park, Yuki Arase, Baotian Hu, Wei Lu, Derry Wijaya, Ayu Purwarianti, and Adila Alfa Krisnadhi, editors, **Findings of the Association for Computational Linguistics: IJCNLP-AACL 2023 (Findings)**, pp. 431–448, Nusa Dua, Bali, November 2023. Association for Computational Linguistics.

[6] Ankan Mullick, Abhilash Nandy, Manav Kapadnis, Sohan Patnaik, Raghav R, and Roshni Kar. An evaluation framework for legal document summarization. In Nicoletta Calzolari, Frédéric Béchet, Philippe Blache, Khalid Choukri, Christopher Cieri, Thierry Declerck, Sara Goggi, Hitoshi Isahara, Bente Maegaard, Joseph Mariani, Hélène Mazo, Jan Odijk, and Stelios Piperidis, editors, **Proceedings of the Thirteenth Language Resources and Evaluation Conference**, pp. 4747–4753, Marseille, France, June 2022. European Language Resources Association.

[7] Tejpalsingh Siledar, Swaroop Nath, Sankara Muddu, Rupasai Rangaraju, Swaprava Nath, Pushpak Bhattacharyya, Suman Banerjee, Amey Patil, Sudhanshu Singh, Muthusamy Chelliah, and Nikesh Garera. One prompt to rule them all: LLMs for opinion summary evaluation. In Lun-Wei Ku, Andre Martins, and Vivek Srikumar, editors, **Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)**, pp. 12119–12134, Bangkok, Thailand, August 2024. Association for Computational Linguistics.

[8] Kazuki Fujii, Taishi Nakamura, Mengsay Loem, Hiroki Iida, Masanari Ohi, Kakeru Hattori, Hirai Shota, Sakae Mizuki, Rio Yokota, and Naoaki Okazaki. Continual pre-training for cross-lingual llm adaptation: Enhancing japanese language capabilities. In **Proceedings of the First Conference on Language Modeling**, COLM, p. (to appear), University of Pennsylvania, USA, October 2024.

[9] Tim Dettmers, Artidoro Pagnoni, Ari Holtzman, and Luke Zettlemoyer. Qlora: Efficient finetuning of quantized llms. In A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine, editors, **Advances in Neural Information Processing Systems**, Vol. 36, pp. 10088–10115. Curran Associates, Inc., 2023.

[10] OpenAI. Gpt-4o-mini: A large language model by openai. https://openai.com/index/gpt-4o-mini-advancing-cost-efficient-intelligence/, 2024. Accessed: 2024-12-21.

[11] Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, brian ichter, Fei Xia, Ed Chi, Quoc V Le, and Denny Zhou. Chain-of-thought prompting elicits reasoning in large language models. In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh, editors, **Advances in Neural Information Processing Systems**, Vol. 35, pp. 24824–24837. Curran Associates, Inc., 2022.

[12] Jacob Cohen. A coefficient of agreement for nominal scales. **Educational and psychological measurement**, Vol. 20, No. 1, pp. 37–46, 1960.

[13] J. Richard Landis and Gary G. Koch. The measurement of observer agreement for categorical data. **Biometrics**, Vol. 33, No. 1, pp. 159–174, 1977.

This work is licensed by the author(s) under CC BY 4.0
(https://creativecommons.org/licenses/by/4.0/).

付録

表 6　判示事項と生成要約文の例

判示事項	追突事故に遭った妊婦が事故の約二か月後に早産で未熟児を出産し、難聴や精神発達遅滞の後遺障害が生じたことについて、事故との相当因果関係を肯定し、その損害の賠償を認めた事例
生成要約文	交通事故に遭った妊婦が早産し、未熟児として出生した子供が高度の難聴及び精神発達遅滞の後遺障害が残った場合につき、事故と障害との間に相当因果関係が認められた事例

表 7　評価器用プロンプト

システムプロンプト	あなたは法律の専門家です。与えられた判決書とその要約文から要約文の評価を JSON 形式で出力してください。\n 評価観点と評価基準はは以下の通りです。\n「事案の概要を含む」\n - スコア 2: 含んでいる \n - スコア 1: 含んでいるが内容が不十分 \n - スコア 0: 含んでいない。もしくは記述されているが判決書本文とは関係がない \n 「争点を含む」\n - スコア 2: 含んでいる \n - スコア 1: 含んでいるが内容が不十分 \n - スコア 0: 含んでいない。もしくは記述されているが判決書本文とは関係がない \n 「争点に対する判断を含む」\n - スコア 2: 含んでいる \n - スコア 1: 含んでいるが内容が不十分 \n - スコア 0: 含んでいない。もしくは記述されているが判決書本文とは関係がない \n「請求に対する結論を含む」\n - スコア 2: 含んでいる \n - スコア 1: 含んでいるが内容が不十分 \n - スコア 0: 含んでいない。もしくは記述されているが判決書本文とは関係がない \n 「判決書に対して矛盾を含まない」\n - スコア 2: 矛盾がない \n - スコア 1: 矛盾ではないが判決書に記述されていない内容を含む \n - スコア 0: 矛盾を含む \n 「文法上適切である」\n - スコア 2: 問題なし \n - スコア 1: 判示事項として不自然な表現を含む（です・ます調など）\n - スコア 0: 文法的に正しくない表現を含む。もしくは文として破綻している \n 「適切な長さである」\n - スコア 2: 適切な長さである \n - スコア 1: 実用に耐えうるが、より簡潔に記述するべきである \n - スコア 0: 実用に耐えない程度に長い \n 「論理的な一貫性がある」\n - スコア 2: 論の展開や記述の順番に破綻がない \n - スコア 1: 実用に耐えうるが、論理の展開や記述の順番が不自然である \n - スコア 0: 論理の展開や記述の順番に破綻がある \n 上記の基準に基づいてスコアを 0,1,2 のいずれかで出力してください。
ユーザープロンプト	次の判決書とその要約文から要約文の評価を JSON 形式で出力してください。辞書のキーは 8 種類の評価基準の名前にしてください。\n 判決書:[[BODY]]\n 要約文:[[SUMMARY]]\n 評価:
システムプロンプト（コメント付き生成）	あなたは法律の専門家です。与えられた判決書とその要約文から要約文の評価を JSON 形式で出力してください。評価観点と評価基準はは以下の通りです。\n [[ルーブリック略]]\n 上記の基準に基づいてスコアを 0,1,2 のいずれかで出力してください。スコアを出力する前に、要約文に対するフィードバックを出力してください。
ユーザープロンプト（コメント付き生成）	次の判決書とその要約文から要約文の評価を JSON 形式で出力してください。辞書のキーは「要約文に対するフィードバック」と 8 種類の評価基準の名前にしてください。\n 判決書:[[BODY]]\n 要約文:[[SUMMARY]]\n 評価:

表 8　ROUGE-1 と「①事案の概要を含む」の人手評価が逆転している例

生成判示事項	ROUGE-1	①事案の概要を含む
原告が，その所有に係る土地と，隣地である被告ら所有に係る土地との間の境界確定を求め，認容された事例	21.28	2
１亡Ａの業務が過重であったとはいえないから，本件疾病の発症についてＸ社に予見可能性がないし，安全配慮義務違反もないとされた例 \n ２本件死亡がＸ社の業務に起因しているとは認められないとされた例	33.09	0

This work is licensed by the author(s) under CC BY 4.0
(https://creativecommons.org/licenses/by/4.0/).