Extraction of Human Relationship Terms Associated with Characters and Speakers from Novel Lines

***Abstract*—Non-task-oriented dialogue systems are increasingly expected to serve as long-term partners, building trusting relationships with users. Achieving human-like, natural dialogues is crucial, as inaccurate responses can discourage users. Previous studies have explored methods such as updating personas, enhancing response quality, and utilizing dialogue history. This study focuses on extracting human relationships from dialogues in novels, which serve as long-term interactions between characters. As a preliminary study on a technique for extracting human relationships from dialogues between a user and a dialogue system, we propose a model for extracting human relational terms from lines from novels. Experimental results show that the proposed model can extract relational terms with an F1-score of 0.6647.**

Keywords— novel, line, relation extraction, human relationship, relational term, dialogue, pre-trained language model

I. Introduction

Non-task-oriented dialogue systems have been increasingly expected to serve not only as short-term dialogue partners, but also as long-term partners that satisfy users' long-term dialogue needs [1]. For non-task-oriented dialogue systems intended for long-term dialogue to build a trusting relationship with the user, achieving human-like and natural dialogue is crucial. However, when the system generates inaccurate responses, users may become confused and feel discouraged from continuing the dialogue [2]. To address this issue, previous studies have explored approaches such as automatically updating personas to maintain dialogue consistency [3], enhancing response quality and accuracy by incorporating specific knowledge [4], and utilizing dialogue history [5].

Especially in long-term dialogues, it is important to extract and utilize useful information from the dialogues to cope with the limited input length of language models. Related studies include approaches that focus on user persona information [5] and those that utilize users' social relationships [6]. Understanding users' social relationships through dialogue is essential for dialogue systems to interpret user intent and generate appropriate responses. However, Liang, et al. [6] focus solely on social relationships among multiple speakers, making it difficult to capture user relationships from one-on-one dialogues between users and systems. Therefore, this research aims to develop a chat-oriented dialogue system that can build a long-term trusting relationship with users by extracting and utilizing user relationships that emerge during the dialogues.

In order to consider methods for extracting and utilizing human relationships from user-system interactions, dialogue data containing such relationships is necessary. However, existing dialogue corpora do not include data that involve personal relationships and long-term dialogues. In this study, we focus on the dialogue between characters in novels. In novels, dialogues between characters occur chronologically as the story progresses. Therefore, the dialogues in novels can be regarded as long-term dialogues between characters. Furthermore, since the various relationships among characters are reflected in their dialogues, the dialogues tend to include human relationships.

Several studies [7, 8] have been conducted to extract character relationships from novel texts. These studies utilize character lines and surrounding narrative descriptions to extract such relationships. However, narrative descriptions cannot be utilized in dialogues between a dialogue system and a user. Additionally, Japanese is a language in which both the subject and the speaker are often omitted.

In this paper, as a preliminary study on a technique for extracting human relationships from dialogues between a user and a dialogue system, we propose a model for extracting human relationship terms associated with characters and speakers from novel lines. Here, the character appearing in a novel line is assumed to be a person who appears in the user's utterance within a dialog system, and the speaker of the line is assumed to be the user in the system.

II. Related Work

Several studies [6, 9] have been conducted to estimate human relationships from dialogues. Liang, et al. [6] proposed a method called "SocAoG" to estimate human relationships and personal attributes among speakers from dialogues between multiple people. In this method, each speaker is first represented as a node, and the various human relationships are represented as edges. Then, each speaker node hierarchically stores the speaker's attribute information, such as gender, age, and occupation. By updating the nodes and edges according to the dialogue, the method can capture changes in human relationships over time. They utilized two datasets, DialogRE and MovieGraph, annotated with human relationships. Experimental results showed that the F1 value was 69.5% for DialogRE and 64.1% for MovieGraph.

Tigunova, et al. [9] proposed "PRIDE", a neural multi-label classifier, based on BERT and Transformer for creating a conversation representation to estimate speakers' relationships from dialogues. PRIDE utilizes dialogue structure and augments it with external knowledge about speaker features and conversation style. Experimental results showed that the F1 value was 0.38 for the dataset FiRe, which labeled the relationships between characters in a movie, significantly outperforming previous baselines.

These studies estimate human relationships among speakers from dialogues between multiple people. On the other hand, our study assumes a long-term one-on-one dialogue between a user and a dialogue system, and extracts the user's human relationships from the user's utterances. In addition, there are very few large-scale Japanese dialogue corpora that include human relationships. Therefore, we focus on lines from novels. Since dialogues in novels include various human relationships, they can serve as an alternative to existing dialogue corpora.

Bick, et al. [10] proposed "PALAVRAS-DIP", a rule-based system for automatic identification of characters and their social profiles in Portuguese and Brazilian literature. PALAVRAS-DIP recognizes the character names and extracts their attributes and family relationships between the characters. Similar to Japanese, Portuguese is a language where the subject is often omitted, therefore, a rule-based method utilizing verb conjugation information and context is used to complement the subject. In the shared task evaluation DIP for historical novels, character identification achieved the average F-score of 63.4%, alias resolution through co-reference analysis achieved 68.1%, and family relationship extraction achieved 15.5%. While the rule-based methods demonstrated stable performance, it was confirmed that extracting information provided in complex contexts, such as family relationships, remains challenging.

We aim to realize rule-based and language model-based extraction methods in order to enable the extraction of human relationships even from complex contexts. In this paper, as a preliminary study, we propose a model for extracting human relationship terms associated with characters and speakers from novel lines. We also focus on lines with omitted subjects and compare the extraction performance of each method with that of lines without subject omission.

III. Comparison of Existing Corpora and Novel Lines

A. Analysis of Diversity of Expressions

In existing daily dialogue corpora [11, 12], utterances containing personal information or information that the speaker does not want to disclose are masked or deleted for privacy protection. In addition, the number of utterances per dialogue is small. The Japanese Daily Dialogue Corpus (JDD) [11] is a representative corpus of Japanese dialogues. However, the average number of utterances per dialogue is only 9.94. The RealPersonaChat (RPC) [12] corpus also has an average of 30.09 utterances per dialogue. Even in everyday conversations between humans, utterances containing human relationships do not appear frequently and tend to emerge over long-term interactions. Therefore, they are not included in the existing dialogue corpora.

Lines from novels tend to include words or expressions that describe the characters' personalities and backgrounds and indicate relationships between characters. They also contain various emotional expressions such as conflict, intimacy, and sympathy. In order to give the characters individuality, the tone of lines is also varied. In systems that engage in long-term dialogue with users, it is necessary to anticipate that a user's tone of dialogue may change depending on their mood that day. Therefore, the dialogue system we aim to develop must be capable of extracting human relationships from dialogues that include a variety of tones and rich emotional expressions. The lines from novels are the best data to be used for considering the method of extracting human relationships. In addition, the dialogues in novels unfold as natural dialogues, without the constraints of specific tasks. These elements are useful for learning ambiguous expressions and context-dependency that occur in real dialogues.

B. Proportion of Utterances with Human Relationships

In order to verify the validity of lines from novels as a dataset for our study, we investigate the extent to which utterances containing human relationships are found in the existing daily dialogue corpora [11, 12] and novel lines. We extract candidate utterances expressing human relationships (called relational sentence candidates) that contain at least one word indicating a person's name and at least one term representing a human relationship (called relational terms) within a single utterance, and investigate their number and proportion.

Lines are extracted from 18 novels set in the real world, published on the "Shōsetsuka ni Narō [1] (Let's Become a Novelist)" to obtain lines equivalent to those in the existing corpora. To extract human names from utterances, we use GiNZA[2]. As relational terms, we utilize 147 terms manually selected from the "Kadokawa Ruigo Shinjiten[3]" and the "Burui Goihyo [4] (Word List by Semantic Principles, Revised and Enlarged Edition)".

Table I shows the results of the investigation. As shown in Table I, the number and percentage of relational sentence candidates in novels are higher than those in the existing corpora, suggesting that lines from novels may contain a variety of human relationships. Moreover, since novels will continue to be published in the future, the quantitative problem can be addressed. Therefore, in this study, we use lines extracted from novels as data.

TABLE I. INVESTIGATION RESULTS

	JDD	RPC	Novels (18)
Number of sentences (utterances or lines)	41,737	41,737	40,515
Relational sentence candidates (%)	248 (0.59)	53 (0.13)	1,298 (3.20)

[1] https://syosetu.com/

[2] https://megagonlabs.github.io/ginza/

[3] https://www.kadokawa.co.jp/product/199999011700/

[4] https://clrd.ninjal.ac.jp/goihyo.html

TABLE II. CATEGORIES OF NOVEL LINES CONTAINING HUMAN RELATIONSHIPS

	Category	Definition	Example of Lines and relationships between a character's name and a relational term
Character-related data	1	One person's name and relational terms are present	「山田は友達だよ (Yamada is my friend.)」[Yamada - friend]
	2	Two persons' names and related terms are present	「結城と和樹は友達なのね (Yuki and Kazuki are friends.)」[Yuki - friend - Kazuki]
	3	Multiple human relationships are present	「私の兄，和樹に恋人がいた (My brother, Kazuki, had a lover.)」 [I - brother - Kazuki], [Kazuki - lover]
	4	Human relationships requiring contextual understanding	「結城さんは友達だったけど,これからは恋人 (Yuki was my friend, but from now on, he is my lover.)」 [Yuki - Lover]
Speaker-related data	5	Only the speaker's human relationships are present	「今日は妻との結婚記念日 (Today is (my) wedding anniversary with (my) wife.)」 [Speaker - wife]
	6	Human relationships between the speaker and characters are present	「妻とショッピング中に，友人の和樹に会った (While shopping with (my) wife, (I) met (my) friend Kazuki.)」 [Speaker - wife] [Kazuki – friend]

IV. MODELS FOR EXTRACTING RELATIONAL TERMS ASSOCIATED WITH CHARACTERS AND SPEAKERS

A. Dataset Overview

"Shōsetsuka ni Narō" is one of the most popular websites for posting original novels in Japan. In this study, we propose a model to extract relational terms associated with characters and speakers using lines from novels set in the real world, published on this website. Sentences enclosed in Japanese quotation marks ("「 」" and "『 』") are considered characters' lines. A dataset is constructed by extracting 2,000 lines from 64 novels containing one or more relational terms. To enable the model to train human relationships that appear in actual dialogues, we selected the top 47 most frequent words in the Japanese Daily Dialogues Corpus from a set of 147 words manually extracted from the Kadokawa Thesaurus and Classified Vocabulary Database.

To confirm the characteristics of lines containing relational terms, we classified such lines into six categories. Table II shows the results of classifying lines containing human relationships. To train a model for extracting relational terms associated with characters and speakers, we further classified the dataset into two types: characters-related data and speaker-related data. As a premise, in Japanese, the subject and first-person possessive (e.g., my) in the speaker's utterances tend to be omitted, as shown in the speaker-related data. Therefore, words in parentheses in the English sentences indicate elements that are omitted in the Japanese sentences.

The model proposed in this paper takes a line and a person's name as input, and predicts the relational terms associated with that name based on the line. For a line containing multiple names, the data is expanded for each name, and positive or negative labels are assigned accordingly. For example, in the case of "Yuki and Kazuki are friends," the data is expanded into two sets, [Yuki - friend] and [Kazuki - friend], ensuring a one-to-one relationship between the character's name and the relational term. Then, the dataset is constructed after adjusting for data imbalance between categories.

TABLE III. DATASET

Data Type	Character				Speaker		Negative example
Category	1	2	3	4	5	6	7
After adjustment	100	100	100	76	188	188	752

B. Category Distribution of Dataset

The dataset consists of a combination of character-related data and speaker-related data, balanced at a 1:1 ratio. Positive and negative examples of each dataset were also adjusted to be uniform, resulting in the data distribution shown in Table III.

To train the extraction model for the relational terms associated with the character, we use 376 positive examples from Categories 1 to 6 and 376 negative examples for characters in Category 7. To train the extraction model for the relational terms associated with the speaker, we use 376 positive examples from Categories 5 and 6, and 376 negative examples for speakers in Category 7.

C. Relational Terms Extraction Models

Using the dataset, we construct a model to extract relational terms associated with characters and speakers from lines and evaluate its extraction performance. For the language model, we use the pre-trained BERT[5] provided by Tohoku NLP Group. To extract relational terms by classifying each token into three categories based on BIO tags, we construct an extraction model by adding a classification layer, consisting of a Linear Layer and a Softmax function, to the final layer of BERT.

To evaluate the extraction performance of characters and speakers, we compare the performance of two models, "BERT (Character)" and "BERT (Speaker)", which were trained separately on the character-related datasets and speaker-related datasets, with the proposed model "BERT (Character + Speaker)". The structure of the proposed model and the individually trained models is the same. The performance of the individually trained model is compared under the same data conditions as the proposed model by using 752 positive and 752 negative examples.

[5] https://huggingface.co/tohoku-nlp/bert-base-japanese-v3

TABLE IV. EXTRACTION PATTERNS BY NISHIHARA, ET AL. [13]

<table>
<tr><th>Extraction Patterns and Examples</th></tr>
<tr><td>P1 の P2&R [結城の兄 (Yuki's brother)]</td></tr>
<tr><td>P1 の R(の|は)P2 [結城の兄の太郎 (Taro, Yuki's brother)]</td></tr>
<tr><td>P2(が|は|も)~P1 の~R [結城は和樹の中学の友達 (Yuki is Kazuki's middle school friend)]</td></tr>
<tr><td>P1(が|は|も)~R の P2 [結城は時々兄の太郎と (Yuki sometimes (does somthing) with his brother, Taro)]</td></tr>
<tr><td>P1(に|には)P2 という R [結城に太郎という兄が (Yuki has a brother named Taro)]</td></tr>
<tr><td>P1 が~R、P2 [結城が友人、和樹と (Yuki and his friend, Kazuaki)]</td></tr>
<tr><td>P1(が|は|も)~P2&R(と|を|に|の) [太郎が公園で妹と (Taro is with his sister at the park)]</td></tr>
<tr><td>P2&R(が|は|も)~P1(と|を|に|の) [兄が結城を ((His) Brother (does something to) Yuki)]</td></tr>
</table>

The language model used in this experiment differs in the way that data is input for characters and speakers. For the character data, the model combines the line and the character's name mentioned in the line as [CLS] line [SEP] character's name (e.g., [CLS] I played with my friend Takumi today [SEP] Takumi). Since the subject is omitted in the speaker-related data, the model substitutes the unseen speaker with the word "speaker". It then predicts the associated relational terms by inputting [CLS] line [SEP] speaker (e.g., [CLS] (I) went on a trip with (my) sister [SEP] speaker).

In this paper, we compare the performance of a rule-based extraction model (Rule), the proposed model "BERT (Character + Speaker)" and models "BERT (Character)" and "BERT (Speaker)", which are trained independently with characters-related data and speaker-related data, respectively.

For the rule-based model, we utilize the pattern proposed by Nishihara, et al. [13]. Table IV shows the extraction patterns. These extraction patterns were manually constructed for the purpose of extracting characters and related terms from novel text. Here, "P" is the character's name, and "R" is the relational term. From a preliminary investigation, we confirmed that in the speaker-related data, relational terms (e.g., "senior" and "teacher") appeared consecutively with character's names more frequently than in the character-related data. Therefore, in this paper, we define "P1R: the consecutive appearance of a character's name and a relational term" as a new pattern for the speaker-related data, and add it to the extraction patterns.

D. Experimental Setup

We compare the performance of a rule-based extraction model (Rule), the proposed model "BERT (Character + Speaker)" and models "BERT (Character)" and "BERT (Speaker)". Each BERT model is fine-tuned as a three-class classification using BIO tags and evaluated using stratified 10-part cross-validation. True Positive, False Positive, and False Negative are counted, and overall Precision, Recall, and F1-score are calculated as evaluation metrics. For each test, the model from the epoch with the minimum test loss is adopted.

Since there are few utterances indicating human relationships in daily dialogues, we focus on recall to identify omissions in each category. In addition, we also evaluate Precision and F1-score.

E. Results

The performance of the Rule model and the BERT (Character + Speaker) model is shown in Table V. The Rule model achieves a Precision of 0.719, higher than that of the BERT model, suggesting that most extracted relational terms are associated with the character's name. On the other hand, the BERT model has a lower Precision of 0.5687. However, a Recall of 0.8024 and an F1 value of 0.6647, indicating a better overall balance. From these results, it can be said that the Rule model shows superior performance on specific patterns, while the BERT model enables the extraction of a wide range of relational terms, due to the diversity of relational terms in the dataset.

Table VI shows each Recall value for the rule-based model (Rule), the BERT (Character) model, the BERT (Speaker) model, and the BERT (Character + Speaker) model. The Rule model has a lower recall than the BERT (Character + Speaker) model across all categories. The BERT (Character) model showed higher recall in Categories 1 and 2, while the BERT (Speaker) model showed higher recall in Categories 5 and 6. Although the proposed BERT (Character + Speaker) model shows some variation in recall in each category, it has a high overall extraction performance. In the proposed model, the performance of Category 1 was improved because this category often contains only one relational term associated with the speaker, making it easier to extract simple human relationships. Because multiple relationships appeared in Category 3, which showed lower performance, may not have been able to select the appropriate relational terms associated with the character's name.

Next, we focus on Precision values for each category, as shown in Table VII. The Rule model is effective in extracting specific relational terms whose extraction patterns appear in the lines, since its performance is high in Categories 1 and 6. The BERT (Speaker) model showed high Precision in Categories 5 and 6. In contrast, the proposed BERT (Character + Speaker) model showed lower performance in Categories 2 and 4 than the individually trained model, however, maintained similar performance in the other categories, especially in Categories 1, 5, and 6. The reason for the lower performance in Category 2 is thought to be that the extraction of relational terms depends on contextual cues of the characters, and the simultaneous provision of speaker-related data that did not include the characters made the detection of relational terms more difficult. In Category 4 as well, contextual understanding is necessary, requiring extensive reference to the surrounding context of the line. In contrast, the speaker-related data tends to use the relational terms directly without needing deep contextual understanding. Therefore, it is considered that the proposed model has difficulty in detecting relational terms in lines that require contextual understanding.

The F1-scores for each category are shown in Table VIII. The Rule model showed lower performance than the BERT model in all categories. The performance of the BERT (Character) model was higher in Categories 1 and 2, while in the BERT (Speaker) model, the F1 values were higher in Categories 5 and 6. The proposed BERT (Character + Speaker)

TABLE V. OVERALL PERFORMANCE

	Precision	Recall.	F1
Rule	**0.719**	0.3	0.424
BERT (Character + Speaker)	0.5687	**0.8024**	**0.6647**

TABLE VI. RECALL VALUES BY CATEGORY OF NOVEL LINES CONTAINING HUMAN RELATIONSHIPS

	1	2	3	4	5	6
Rule	0.3247	0.3116	0.2866	0.3167	0.125	0.1333
BERT (Character)	0.8084	**0.842**	**0.6367**	**0.7176**	-	-
BERT (Speaker)	-	-	-	-	0.8604	**0.9129**
BERT (Character + Speaker)	**0.8671**	0.7564	0.6084	0.674	**0.8722**	0.887

TABLE VII. PRECISION VALUES BY CATEGORY

	1	2	3	4	5	6
Rule	0.8289	**0.8267**	0.5165	**0.7600**	0.6000	**1.0000**
BERT (Character)	0.7559	0.7929	**0.6390**	0.6989	-	-
BERT (Speaker)	-	-	-	-	**0.8806**	0.8850
BERT (Character + Speaker)	**0.8416**	0.6787	0.6151	0.5632	0.8715	0.8424

TABLE VIII. F1-SCORE BY CATEGORY

	1	2	3	4	5	6
Rule	0.4666	0.4526	0.3686	0.4471	0.2069	0.2352
BERT (Character)	0.7813	**0.8167**	**0.6378**	**0.7081**	-	-
BERT (Speaker)	-	-	-	-	0.8704	**0.8987**
BERT (Character + Speaker)	**0.8542**	0.7154	0.6117	0.6136	**0.8718**	0.8641

model showed lower performance in Categories 2 and 4 compared to the individually trained BERT model. However, the performance in the other categories remained at or above the same level. In particular, the best F1-score was achieved in Categories 1 and 5, and high extraction performance was achieved for lines containing simple human relationships. We also confirmed that Category 3, where multiple relationships exist in the same line, and Category 4, which requires contextual understanding, are more difficult to extract because they require deeper contextual comprehension.

From the above results, we confirmed that the proposed BERT (Character + Speaker) model can extract various relational terms associated with characters and speakers without a significant performance drop compared to the BERT (Character) and the BERT (Speaker) models, although there are issues with Categories 3 and 4, which require deep contextual understanding.

F. *Error Analysis*

It is considered that different factors may affect the incorrect extraction of relational terms by the proposed model "BERT (Character + Speaker)" for each category. Therefore, we analyze errors by category.

First, Categories 1, 2, and 5 have only one relational term candidate, so there is little possibility of incorrect extraction. In addition, the extraction performance tends to be high for cases in which a character's name and a relational term appear close together, such as "Yamada's mother." On the other hand, incorrect extraction occurs in cases where the context within the lines is complex. In the case of "I am Takashi's lover, not his mother," "lover" is the correct relational term for "I". However, "mother" appearing near the character's name was incorrectly extracted.

Category 3 makes it difficult to select the relational terms to be extracted because multiple relationships exist in the line. For example, in "I'm sorry, but the Kannushi wife's health is more important, so please take care of her. ...Really, Haruharu's father is doing everything,[6]", the relational term associated with "Kannushi" is "wife". However, "father" was extracted instead. Such incorrect extractions suggest the presence of noise caused by multiple relational terms in the lines.

Category 4 causes incorrect extraction in lines where contextual understanding is necessary for extracting relational terms. As an example, in "I and Rinka are married in a net game. Then, one day, we became acquaintances in real life, and then ... Rinka told me "We are married in real life, too."[7]". The relational term "acquaintances", which is associated with "Rinka", should have been extracted. However, "marriage," which is not an actual relationship, was incorrectly extracted.

In Category 6, the presence of a human relationship between the speaker and characters tends to increase false positives. For example, "(My) father's condition seems dangerous. (I'm) going to the hospital now. Comrade Yoshio[8]", in addition to "father" which is a relational term associated with the speaker, "comrade," which has a human relationship with the character, was also incorrectly extracted. The reason for this is likely due to the contextual entanglement of relationships between the speaker and the character.

One common error across all categories was that only a part of the relational terms was extracted. For example, only "先" was extracted from the relational term "先輩." This is thought to be due to insufficient training data.

The above error analysis suggests that the proposed BERT (Character + Speaker) model can extract a variety of relational terms by integrating the input of the character and speaker. However, it is susceptible to noise when the context is complex or when multiple relational terms exist. In the future, it is

[6] Tenri Kurokawa, "Soshite Shoujo wa Akujo no Karada wo Te ni Ireru", Shōsetsuka ni Narō.

[7] Aboon, "Netoge no Yome ga Ninki Aidoru datta Ken ~Kuuru-kei no Kanojo wa Genjitsu demo Yome no Tsumori de Iru~," Shōsetsuka ni Narō.

[8] Bekio, "Reiko no Fuugi ~Akuyaku Reijou to Yobaremashita ga, Tada no Binbou Musume desu~," Shōsetsuka ni Narō.

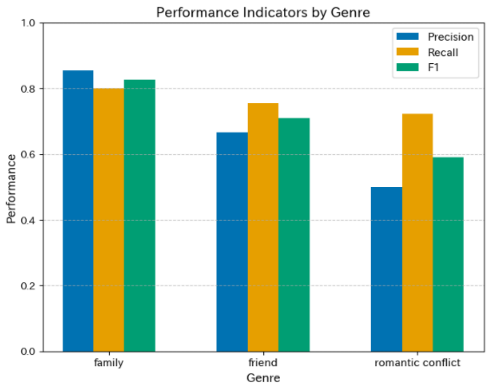

Fig. 1. Performance by Relational Term Genre

necessary to expand the training data and improve the data input method according to the characteristics of each category.

V. Evaluation by Relational Term Genre

In the previous chapter, we evaluated the extraction performance of the proposed BERT (Character + Speaker) model by comparing it with the rule-based model and individually trained models. We focus on the genres of the 47 relational terms in the training data, and classify them into three genres: "Family Relations", "Friend Relations" and "Romantic Conflict Relations" and then evaluate the extraction performance in each genre. Examples of the classified relational terms are shown below.

1. Family Relations: Mother, Father, Brother, Sister
2. Friend Relations: Friends, associates, acquaintances
3. Romantic Conflict Relations: Lovers, Marriage, Enemies

Fig. 1 shows the average Precision, Recall, and F1-score for each relational term genre for the BERT (Character + Speaker) model in the cross-validation. Since the Japanese Daily Dialogues Corpus used as a reference when constructing the training data contained the largest number of relational terms in Family Relations, the performance in Family Relations was higher than those in other genres. This confirms that Family Relations, which tend to appear frequently in daily dialogue, can be extracted with high performance, suggesting that lines from novels may serve as a substitute for daily dialogue. On the other hand, Friend Relations and Romantic Conflict Relations are less likely to appear in daily dialogue, and the limited amount of training data may have led to lower extraction performance.

From these results, we confirmed that the proposed model retains the bias inherent in the training data. In the future, it will be necessary to adjust the amount of training data for each genre.

VI. Conclusion

In this paper, we propose a model for extracting relational terms associated with characters and speakers using lines from novels. In response to the problem that Japanese dialogue tends to omit the subject (i.e., the speaker), we proposed a method for simultaneously extracting relational terms associated with characters and speakers by devising the data input method.

Experimental results showed that the proposed method can extract a variety of relational terms associated with characters and speakers without significant performance degradation compared to a model in which characters and speakers are trained separately. In addition, we classified the relational terms into three genres and evaluated the extraction performance for each genre. As a result, we confirmed that the extraction performance of Family Relationships, which is most likely to appear in daily dialogue, was the highest among all genres. This suggests that lines from novels may be a substitute for daily dialogues as data for training to extract relational terms. On the other hand, the extraction performance of Romantic Conflict Relations was low, confirming the bias in the model caused by the training data.

In the future, we will consider how to input data according to the characteristics of each category and how to increase the size of training data. Then, we will consider a method for extracting human relationships based on relational terms and utilize them in a chat-oriented dialogue system.

VII. References

[1] J. Xu, A. Szlam, and J.Weston, "Beyond Goldfish Memory: Long-Term Open-Domain Conversation," In Proc. of the the 60th Annual Meeting of the Association for Computational Linguistics, pp.5180-5197, 2022.

[2] B. Martinovsky and D. Traum, "The Error Is the Clue: Breakdown in Human-Machine Interaction," In Proc. of Error Handling in Spoken Dialogue Systems, pp.11-16, 2003.

[3] X. Xu, Z. Gou, W. Wu, Z.Y. Niu, H. Wu, et al., "Long Time No See! Open-Domain Conversation with Long-Term Persona Memory," In Findings of the Association for Computational Linguistics: ACL 2022, pp.2639-2650, 2022.

[4] Y. Gao, Y. Xiong, X. Gao, K. Jia, J. Pan, et al., " Retrieval-Augmented Generation for Large Language Models: A Survey," arXiv preprint arXiv:2312.10997, 2023.

[5] S. Han, S. J. Park, C. W. Kim and Y. M. Ro, "Persona Extraction Through Semantic Similarity for Emotional Support Conversation Generation," In Proc. of 2024 IEEE International Conference on Acoustics, Speech and Signal Processing, pp.11321-11325, 2024.

[6] L. Qiu, Y. Liang, Y. Zhao, P. Lu1, B. Peng, et al., "SocAoG: Incremental Graph Parsing for Social Relation Inference in Dialogues," In Proc. of the 59th Annual Meeting of the Association for Computational Linguistics, pp.658-670, 2021.

[7] C. Fan, and Y Li, "Network Extraction and Analysis of Character Relationships in Chinese Literary Works," Internatinal Journal of Computational Intelligence and Neuroscience, 2022.

[8] Kim, Evgeny and Roman Klinger, "Frowning Frodo, Wincing Leia, and a Seriously Great Friendship: Learning to Classify Emotional Relationships of Fictional Characters," arXiv preprint arXiv:1903.12453, 2019.

[9] A. Tigunova, P. Mirza, A. Yates, and G. Weikum, "PRIDE: Predicting Relationships in Conversations," In Proc. of the 2021 Conference on Empirical Methods in Natural Language Processing, pp.4636-4650, 2021.

[10] E. Bick, "Extraction of Literary Character Information in Portuguese," Linguamática, Vol.15, No.1, pp.31-40, 2023.

[11] R. Akama, J. Isobe, J. Suzuki, and K. Inui, "Construction of a Japanese Daily Dialogue corpus," In Proc. of the 29th Annual Meeting of the Association for Natural LanguageProcessing, pp.108-113. 2023. (in Japanese)

[12] S. Yamashita, K. Inoue, A. Guo, S. Mochizuki, T. Kawahara, and R. Higashinaka, "RealPersonaChat: A Realistic Persona Chat Corpus with Interlocutors' Own Personalities," In Proc. of the 37th Pacific Asia Conference on Language Information and Computation, pp.852-861, 2023.

[13] H. Nishihara and K. Shirai, "Relation Extraction of Characters from Narrative Texts," In Proc. of the 21st Annual Conference of the Association for Natural Language Processing, pp.628-631, 2015.