

Annotating a Driving Experience Corpus with Behavior and Subjectivity

Ritsuko Iwai^{1,2}, Daisuke Kawahara¹, Takatsune Kumada^{1,2}, Sadao Kurohashi¹

¹Graduate School of Informatics, Kyoto University

Yoshida-honmachi, Sakyo-ku, Kyoto, 606-8501, Japan

²RIKEN BSI TOYOTA Collaboration Center, RIKEN

Hirosawa, Wako-shi, Saitama, 351-198, Japan

ritsuko.iwai@riken.jp, {dk, t.kumada, kuro}@i.kyoto-u.ac.jp

Abstract

Personality is the important internal framework that we use when we communicate with the others; thus, it is critical for vehicle-driver communication. Evaluating an individual's personality requires associating behavior and subjectivity. To develop the vehicle-driver communication system that will be most applicable to actual driving settings, a corpus is necessary that includes colloquial and daily experiential expressions along with their emotions, polarity, and sentiments. In addition, the corpus must include a wide range of subjective measures, such as human judgements, perceptions, and cognitions during car driving. Thus, we construct a driving experience corpus (DEC) that constitutes 253 blog articles (7,831 sentences) with the following four manually annotated tags: (1) driving experience (DE), (2) other's behavior (OB), (3) self-behavior (SB) and (4) subjectivity (SJ). In this paper, we describe the guidelines, corpus specification and agreement analysis between annotators. We identified three difficulties: the extended self, important information, and voice in mind. We conducted automatic annotation experiment on the corpus using Conditional Random Fields (CRF). The results indicated F-Scores of .768, .478, .534, and .749 for DE, OB, SB, and SJ, respectively, on the test set. Our error analysis reveals difficulties in

interpreting nominatives and recognizing facts.

1 Introduction

Given the drastic improvements in the field of machine learning, agents and robots, systems share or play a role in tasks that used to be fully controlled by humans. Many companies are, for example, competing to realize automatic driving and are conducting test driving on roads. However, it is critical that systems understand individual human differences and predict their behavior, especially under the condition that a miss may bring about a serious fatal accident. Systems must catch the attention of the driver and make him take the appropriate action. Thus, it is important for systems to communicate with their users in accordance with their individual differences. Tapus et al. (2008), for example, demonstrated that participants were more engaged in rehabilitation exercises with robots that behaved and gave messages similarly to the personalities of the participants. We regard it as important that a system communicates with its user via a language based on his/her personality to bridge the inconsistencies between the system's requirements and the user's attention and actions.

A personality is defined as an "abstraction used to describe and explain the coherent patterning over time and space of affects, cognitions, desires and the resulting behaviors that an individual experiences and expresses" (Revelle & Condon, 2015, p. 70). We consciously or unconsciously evaluate others' personalities, predict their next behaviors or utterances and sometimes adjust our communication style as a result. Thus,

understanding the human personality and related behaviors is quite critical for this system’s implementation. The definition of a personality suggests two important things that a human or a system needs to understand: behaviors and their psychological reactions. In other words, behavior and subjectivity are the critical determinants.

A corpus designed for this goal is more critical than anything else, and it must include the following three characteristics:

- Colloquial and daily expressions under actual driving experiences,
- Subjectivity (opinions, perceptions, cogitations, thinking, and emotions),
- Driving behaviors.

Thus, for the first step, we construct a Driving Experience Corpus (DEC) in Japanese, which consists of 253 blog articles (7,831 sentences). The following two sentences indicate examples from the actual DEC.

- (1) Subjectivity ちょうどタイミングよくあと一歩のところだし Others_Behavior 踏切が開いたら車がダッシュ。
[Subjectivity Just one more step closer to catch up], [Others_Behavior the railroad crossing was opened and the car dashed].
- (2) Others_Behavior そのさらに先の信号でまた止まったのでまた Self_Behavior 淡々と追いかけた。
[Others_Behavior Since the next signal turned to red and the car stopped again], [Self_Behavior I continued chasing the car again].

Figure 1: Example of the annotation

In this study, we report the process of corpus construction (Figure 2), including the problems

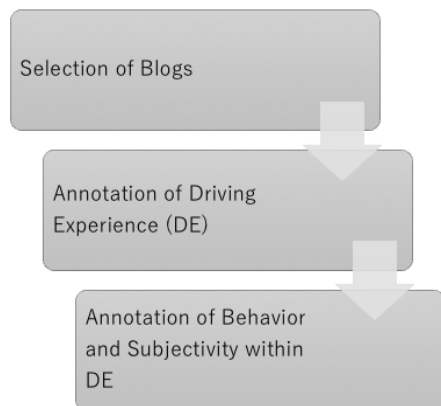


Figure 2: Process of annotation

that occurred and their solutions, the evaluation of the annotation, and an experiment to automatically tag a text using a Conditional Random Fields-based model trained on the DEC.

2 Related Works

In this section, we briefly introduce the related works on personality, driving and annotations.

2.1 Personality

Psychological studies have suggested that personality is the universal framework (e.g., Costa and McCrae, 1992; Goldberg, 1992; Goldberg et al., 2006; McCrae et al., 1996; McCrae et al., 2005; Saucier & Goldberg, 2001).

In the field of Natural Language Processing, there are several studies that predict the personalities of authors from SNS (Social Network Services) texts both in English (e.g., Golbeck et al., 2011; Park et al., 2015; Plank & Dirky, 2011; Schwartz et al., 2013) and in Japanese (e.g., Nasukawa et al., 2016; Nasukawa & Kamijo, 2017). These studies, however, focus on the detection of authors’ personalities, while our goal is to associate behaviors and personalities.

2.2 Traffic, Transportation and Driving

Previous studies related to driving in Japan basically focused only on driving actions, traffic rules and ontologies (e.g., Takayama et al., 2017; Kawabe et al., 2015; Suzuki et al., 2015; Taira et al., 2014). These datasets, however, are not publicly available. Furthermore, human-centered or human experiential foci are lacking.

2.3 Annotation

Regarding subjectivity, subjective analysis is one important and active domain. Subjectivity in this analysis refers to “information about any attitudes, beliefs, emotions, opinions, evaluations and sentiment expressed in texts” (Stoyanov & Cardie, 2008, p. 817). The previous studies of subjectivity analysis include emotions (e.g., Laskowski & Burger, 2006; (Tokuhsa et al., 2008, 2009), opinions (e.g., Jakob & Guvebychi, 2008), and sentiments (e.g., Hashimoto et al., 2011; Nakazawa et al., 2018).

Regarding annotated corpora in this domain, four Japanese corpora have been constructed: (1) the University of Tsukuba Corpus Tagged with

Tags	Definitions
Driving Experience (DE)	A car driving experience refers to an experience containing one or two of the following and it specifies the range of this experience: <ol style="list-style-type: none"> 1. Blog experiences including descriptions and impressions of the author's own driving scene, and 2. the blog authors themselves do not drive a car, but actually describe the scenes of other people driving a car and the experiences including the impressions of the scene.
Others Behavior (OB)	This tag refers to the actions, behaviors, and objects that are to be manipulated by humans who are not the authors.
Self-Behavior (SB)	This tag refers to the actions, behaviors and objects that are to be manipulated by the authors, including the present and the past, the date, the time, and the place.
Subjectivity (SJ)	This tag refers to the evaluation of the behaviors. It refers to the emotions, cognitions, thoughts, judgments, and predictions held by the author as a result of actions.

Table 1: Definitions of the tags

Polarity Evaluation (polarity per sentence)², (2) the ACP corpus, (3) the Kyoto University and NTT blog corpus and (4) the Phrase Polarity Corpus. The Tsukuba Corpus consists of 4,309 sentences in the travel services domain with the polarity per sentence¹. The ACP corpus was automatically constructed (Kaji & Kitsuregawa, 2006). This corpus is large (126,610 sentences), but it was not filtered by humans. Hashimoto et al. (2011) constructed the Kyoto University and NTT Blog Corpus of 249 blog articles (4,186 sentences) with sentiment information. These blogs were written by 81 university students in the following four themes: sight-seeing in Kyoto, cellphones, sports or gourmet food. Nakazawa et al. (2018) constructed a phrase-sentiment review corpus (59,758 phrases) with all manually annotated polarity information from the Tsukuba Corpus. These corpora do not focus on driving, thinking and cognition.

Our corpus is different from the existing ones in the following three perspectives:

- It includes colloquial and daily expressions.
- It has a wide range of subjectivity, including emotions, polarity, sentiments,

human judgements, perceptions, and cognitions.

- It covers driving experiences.

3 Guidelines

The guidelines were constructed by an annotation team. The members consisted of the first author, three experienced annotators, and the second author as a supervisor. With 11 driving experiential blog articles, we repeated the annotations, discussions, and revisions of the guidelines.

3.1 Definition of Tags

In this project, we focus on behaviors and subjectivity, especially when driving a car. To this end, we first tag a car driving experience (see the DE example) and annotate the experiences with the behavior and subjectivity tags. The detailed definitions are presented in Table 1.

3.1.1 Driving Experience Tag

To exclude texts that are unrelated to car driving experiences, such as events or opinions, we prepared the DE tag.

² <http://www.nlp.mibel.cs.tsukuba.ac.jp/~inui/SA/corpus/readme-euc.txt>

³ <https://taku910.github.io/crfpp>

3.1.2 Behavior and Subjectivity Tags

The behavior and subjectivity tags are attached only within each DE range. In our guidelines, "behavior" refers to an observable behavior. People who are present in the same scene can "hear", "listen to", "see", or "look" in the same way as others do, and so it is considered that they can be recognized as in common. Behaviors are divided into two tags according to who does the behavior, the author or others, because the delineation is very important. The system needs to differentiate its user and others both in terms of safety and communication. The accidental risk may be notably high if the system fails to deliver the message accurately.

Meanwhile, psychological reactions or mental states such as thinking and feeling are all regarded as subjectivity. The behavior and subjectivity tags are basically annotated within one sentence. All the tagged units, however, change from a word to a clause or to phrases, depending purely on semantics (Figure 3). The subsequent sentences indicate provide an example of one sentence with different units of annotation.

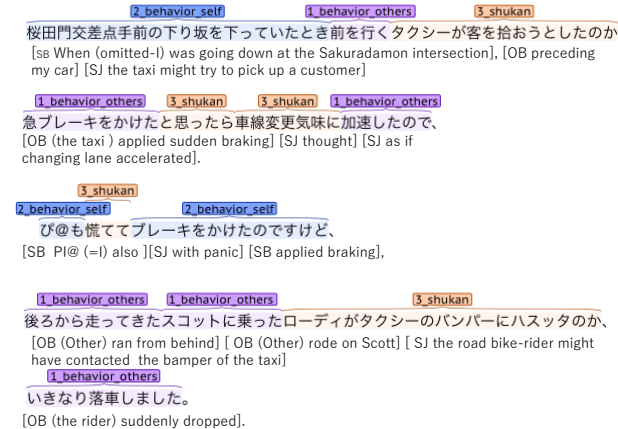


Figure 3: Example of SB, SJ, and OB annotations

3.2 Difficult Cases

In the discussion process, we found three difficulties: (1) the extended self or we-ness, (2) important information, and (3) voice in mind.

Extended self or we-ness: Different from usual situations, it is sometimes quite difficult to differentiate “me” and “you/ other people” when there are multiple people within a car.

In the following sentences, the author was NOT driving the car. She was just sitting at the passenger seat in the car while her husband was driving.

今朝子供たちを幼稚園に送った後、
夫の運転で走っていました。
After sending our children to kindergarten this morning, my husband was driving .

...

青になり、交差点を渡ったら・・・突然
シグナルもださず、・・・、私たちの
走ってるレーンに車線変更してきて、左
前方こすられました。

When we crossed over the intersection, ... a
car suddenly did not send a signal, ... ,
changed the lane where we were running and
scraped the left front of our car.

私、助手席座っていたのですが、キャー
と思わず叫んでた。

I, sitting at the passenger seat, was screaming in
spite of myself.

Although she did not have any control over the car, she was placed into the trouble due to no choice or fault of her own. Thus, we regarded this extended self as self-behavior. On the other hand, “screaming” in the last sentence is only what the author did, which is as originally defined as self-behavior.

Important information: There were many cases in which the annotators were at a loss on or they hesitated to ignore several important information related to driving behaviors. Typically, hesitation results when times and weather conditions appear in the very beginning of the experiences far from the driving behavior. We know they often affect driving behaviors.

Another typical example is important incidents or information that affects the preceding or subsequent tags, even though such information, of course, appears in very close sentences.

さらに、水温計がみるみる上昇していた。
The water temperature gauge was climbing up even more.

60キロくらいまで落ちたところで、
ハザードをつけて路肩に寄せるしかなかった。
When speed fell to about 60 km/hour, I had to put a hazard and pull over.

The water temperature gauge increasing is just a fact and not a behavior or subjectivity. However, it specifically led the author to have to put on their hazard lights and stop the car.

We solved this issue only when this important information that affects behaviors or subjectivity occurs just before or after the main tag. This very close information is easily connectable with the main behavior and subjectivity while the distant

information is difficult to determine to what extent it influences the behavior and subjectivity. For the example sentences, we annotate the first sentence about water temperature gauge with the same tag as the following sentence, which is SB.

Voice in mind: The texts in parentheses (utterances) were mostly included only when they have nominatives and/or verbs as behaviors. The first “embarrassing...” in the example sentence was not annotated with any tags. However, these expressions often appear in the blog writings, and so we annotated such expressions with the subjectivity tag as a voice in mind.

今ではマジェ & スポスタでタラタラとツーリングを楽しんでいるmjですがガキの頃には当時流行のレーサーレプリカで、峠がって峠を攻めたものでした (お恥ずかしい...)

Now mj/the author enjoys touring with Maje/Majesty & Sposuta/Sportster but I used be a little boy going round mountain passes at high speed with a stylish racer replica at the time (embarrassing ...)

4 Annotation

4.1 Corpus Annotation

Manual annotation was performed using the brat annotation tool (Stenetorp et al., 2012) by three annotators. We prepared the driving-related corpus using 10 million blog articles that were filtered using driving related words. After selecting 250 driving related blog articles out of the driving related corpus, the annotators started the actual annotation. For the first round, the 250 articles were divided into three piles and each annotator completed their parts. For the second round, the second annotator reviewed the annotated texts and corrected annotation errors if necessary. If there was a disagreement, they were discussed. These processes resulted in a total of 261 annotated articles (11 of them were those used for making the guidelines).

4.2 Statistics

Table 2 summarizes the DEC and the annotated tags. Although we prepared 261 articles, we only used 253 articles due to parsing errors. The mean of the DE tags is 1.2, at least one or sometimes more in each blogs. . The DE tags range different events and their experiences. Some blogs include more experiences. The SJ tags were the most among behaviors and subjectivity. Driving related

blog articles are, thus, considered as very subjective.

	Total	Means	SD
Sentences	7,831	31.0	21.3
Words	116,136	459.0	336.1
Content words	61,575	243.4	175.9
DE tags	301	1.2	0.5
OB tags	1,873	7.4	6.3
SB tags	2,025	8.0	6.4
SJ tags	4,001	15.8	10.0

Table 2: Corpus statistics (N=253 articles)

4.3 Inter-Annotator Agreement

To evaluate the guidelines, two annotators annotated 15 articles each. The Precision, Recall, and F-scores were calculated for all the tags only when all the DEs were matched (9 articles) to the following: (1) DE, (2) OB, (3) SB, and (4) SJ (Table 3). We used the tags that were annotated by the two annotators described in the previous section as the correct data to evaluate the agreement. The results indicate high agreement. Thus, the guidelines are properly constructed to allow for consistent annotations.

Condition	Tag	P	R	F
Annotator 1	DE	.900	.908	.904
Annotator 2	DE	.965	.899	.931
All	DE	.929	.904	.916
Matched	DE	1.000	1.000	1.000
	OB	.939	.788	.857
	SB	.831	.927	.876
	SJ	.838	.931	.882

Table 3: Annotation agreement evaluations

5 Automatic Annotation Experiments

5.1 Classification

We used Conditional Random Fields (CRF) (Lafferty et al., 2001) to conduct the automatic annotation experiments using our annotated corpus. For these experiments, CRF++ version 0.561³ was used. CRF++ is a sequential tagger that requires a

³ <https://taku910.github.io/crfpp>

template file that specifies the combinations of features.

5.2 Feature Representations

The basic features that are fed to the CRF are the sets of word identities and parts of speech (POSS). In addition to these basic features, we prepared the following five features.

Dependency Relations: Dependency relations are critical information in Japanese such that they were also included in the process. If a word is not the last one within its phrase, it is given its following word. If a word is the last one in its belonging phrase, it is given the content word in its parent phrase (Figure 4).

	Dep.
私/I	は
は/wa-particle	今日
今日/today	、
、	運転
車/car	を
を/wo-particle	運転
運転/drive	し
し/light-verb	た
た/past tense	NULL

Figure 4: An example of the dependency feature

Driving Behavior Words (DBW): We collected the verbs that were derived from the writings about driving experiences by those who participated in psychological experiments at two national research institutes (399 drivers and 250 older drivers over 60 years old). The words include, for example, 運転 ‘drive’, 右折 ‘turning right’, 左折 ‘turning left’, カーブ ‘curve’, 車 ‘vehicle’, 注意 ‘attention’, and 苛々 ‘irritation’.

Driving Behavior Phrases (DBP): Using the DBWs, we extracted the driving behavior predicate phrases from three resources: (1) the same writings as mentioned above, (2) two driving license textbooks for rules and driving (Toyota Nagoya Education Center Inc., 1998, 2004) and (3) driving related blogs. For (1) and (2), we collected all of the predicates using predicate-argument structure analysis with the KNP². For (3), we collected the predicates whose predicate or cases include DBWs. This approach resulted in 7,931 total phrases after the human evaluation. The list includes, for example, ハンドルを切る ‘turn a steering wheel’ and ブレーキを踏む ‘step on the brake.’

Personality words (1): We extended 109 personality related words (Iwai et al., 2017) to 487 words using the word2vec model that is trained with 200 million web-crawled Japanese sentences. These personality extended words are flagged only when zero anaphora resolution results indicated that the nominatives are human when using the Japanese dependency and case structure analyzer KNP⁴.

Personality words (2): We extended 142 Japanese personality related words (Iwai et al., 2017, 2018) to 665 words using the same word2vec model as Personality (1). 395 words were selected based on human evaluations, and they express more direct personality traits such as 心配 ‘anxious’ and 同調 ‘agreeable.’ Conversely, Personality (1) includes more suggestive ones, such as 文化 ‘culture’ and 美術 ‘arts.’ These words are flagged only when zero anaphora resolution results indicated that the nominatives are human when using KNP.

5.3 Experimental Setup

We divided the dataset of 253 articles into training, development and test data and conducted experiments. First, we used all of the features together for the training. After that step, we determined the parameters using the development data. Next, we applied the best model to the test data for the final evaluation.

5.4 Parameters

We adjusted two major parameters, the regularization algorithm and the hyperparameter C, to find the best model. Using the development data, we tested all the combinations of the two parameters (regularization algorithm: L1/L2 and hyperparameter: 0.001/0.01/0.1/1/10). The cutoff threshold of features was not adjusted because our datasets are not very large. We compared the performances and chose L1 and set C to 1.

5.5 Results and Discussion

In our first experiment, we compared the F-scores of each tag using the development data while increasing the training data step by step from 50 articles to 209 articles (Figure 5). For DE and SJ, the F-scores are notably good with 50 articles

⁴ <http://nlp.ist.i.kyoto-u.ac.jp/index.php?KNP>

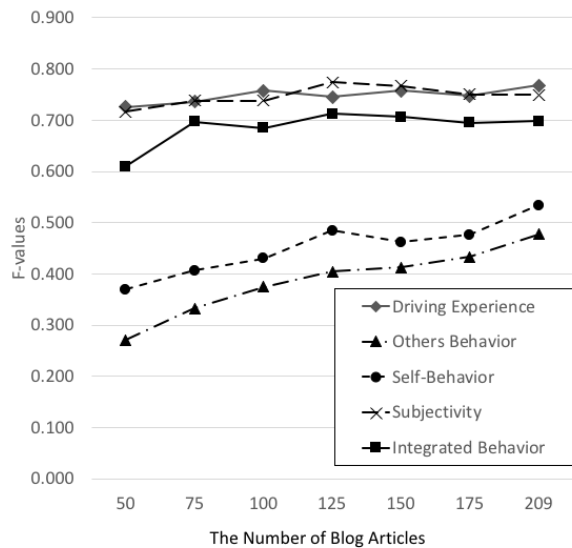


Figure 5: Learning curves of the development data

	DE	OB	SB	SJ	BH
Development	.807	.478	.539	.752	.635
Test	.768	.478	.534	.749	.698

Table 4: F scores of the development and test data

(.760 for DE and .704 for SJ), and the scores increased gradually (.807 for DE and 841 for well

SJ). For the behavioral tags, the F-scores improved from .343 to .478 for OB and from .433 to .539 for SB. We integrated OBs and SBs into one tag, integrated behavior (BH), and BH performed much better than OB and SB, respectively (from .543 to .635). This suggests that the CRF model has learned the behavioral texts but finds it difficult to achieve the correct nominatives.

We applied the model to the test data, and Table 4 lists the performances on the development and test data. The scores of the development data are the same as presented in the results using the 209 articles in Figure 5. The test performance also indicated similar learning patterns to the development data. DE, SJ and BH (DE=.768, SJ=.749, and BH=.698) performed well compared with SB and OB.

Error Analysis. The error analysis on the development data revealed the following two interesting aspects:

- Interpretation of the nominatives (IN), and
- Factuality failure (FF).

Table 5 indicates two examples for each error. The underlined italicized excerpts are the errors that we focus on. The “human” column shows the results of the human annotation and the “CRF” column shows the automatic tags.

	Examples	Human	CRF
IN1	車が僕の真横をかすめた瞬間に僕は急ブレーキで停止。 ‘At the moment when <u>the car passed so close just beside me</u> , I suddenly braked and stopped.’	OB	SB
IN2	待っているうちに前の車が動き出して、私の車が入れるスペースが空いたから、そろそろと前に出た。 ‘While waiting, <u>the preceding car started to move</u> and it formed a space where my car could be, so I moved my car forward gradually.’	OB	SB
FF1	「ここは危険だ」とか「この先はどうなってんだ?」とか思うような道ではとにかく一時停止して、安全を確認してから前進だ。 ‘ <u>On the road</u> where (I/we) think “it is dangerous here” or “How is the road ahead”, <u>(I/we) stop the vehicle and move forward after confirming our safety.</u> ’	SJ	SB
FF2	車が僕の真横をかすめた瞬間に僕は急ブレーキで停止。 At the moment when the car passed so close just beside me, I suddenly braked and stopped.、急停止できたのも奇跡的だった。 ..., <u>the sudden stop</u> was miraculous.	SB	SJ

Table 5: Examples of errors

Both errors result because the system does not have the same ability as a human to read a sentence and comprehend its use, context and knowledge. When humans see and understand any driving behavior, humans regard them as extended others or self-behaviors. In the examples of IN1 and IN2, the CRF model classified the cars as human agents but did not correctly classify whether self or other behaviors could be attributed, although both texts include “I” information.

Another type of error involves the factuality evaluation in each experience. Human annotators read and understand texts with other information as a whole in the experiences and situations, while the CRFs measured the literal semantics. In FF1, ‘(I/ we) stop the vehicle and move forward after confirming our safety’ seems to be a very typical writing that refers to an SB. In this instance, however, this writing provokes the author and readers into being cautious because he/she was involved in a car accident. Thus, the annotators tagged “SJ.” FF2 is the opposite case of the SB tagged as an SJ. The expression “The sudden stop was miraculous.” seems quite subjective if this sudden stop actually has not factually occurred. The author actually bluffed his/her way out of the car accident. Both cases require the system to use the preceding information.

6 Conclusions and Future Work

We constructed a Driving Experience Corpus by annotating behaviors and subjectivity. With the tagged textual data, we constructed a model that sought to understand driving scenes, behaviors as actors or observers and their subjectivity as humans recognize them, which helps the system to communicate with its users in a more colloquial and human-like manner. The annotation scheme and guidelines reflect the ideas and knowledge from psychology. High interannotator agreement shows that the guidelines are applied consistently.

The discrepancies in the process of constructing the guidelines are quite essential characteristics in Japanese blogs or SNS texts. Thus, our guidelines are also useful to those who prepare blogs and SNS texts in Japanese.

Moreover, it is important that the DEC targets driving experiences in Japanese and in any other language. The accurate interpretation of nominatives and factuality recognition are highly

important in the context of the driving environment that entails accidental risks that sometimes lead to injuries and/or death. Understanding drivers’ behaviors and subjectivity helps the system to evaluate their personality. Accurate personality evaluation enhances people’s feelings of being understood (Oishi et al., 2012). Drivers’ feelings of being understood by the system enhance the reliability and trustworthiness of the system, and improve the likelihood of following the system’s messages or cautions.

For our future work, we are currently working on (1) expanding the corpus size with human annotations, (2) conducting further CRF experiments with more features such as modality, polarity, pronouns, zero-anaphora and emotions, and (3) conducting experiments with more recent machine learning methods such as Bidirectional LSTM-CRF (e.g., Huang, Xu, & Yu, 2015; Ma & Hovy, 2016; Reimers & Gurevych, 2017). The errors suggest that simple sequential tagging is not good enough to represent contexts that are located a far distance away.

Acknowledgments

We thank Manami Ishikawa, Marika Horiuchi and Natsuki Nikaido for their careful annotation.

References

- Paul T. Costa and Robert R. McCrae. 1992. Manual for the revised NEO personality inventory (NEO-PI-R) and NEO five-factor inventory (NEO-FFI). Psychological Assessment Resources.
- Jennifer Golbeck, Cristina Robles, Michon Edmondson, and Karen Turner. 2011. Predicting Personality from Twitter. In Proceedings of Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing, 2011 IEEE Third International Conference, pages 149-156.
- Lewis. R. Goldberg. 1992. The development of markers for the Big-Five factor structure. *Psychological Assessment*, 4(1), 26-42. <http://doi.org/10.1037//1040-3590.4.1.26>
- Lewis. R. Goldberg, J.A. Johnson, H. W. Eber, R. Hogan, M.C. Ashton, C.R. Cloninger, and H. G. Gough. 2006. The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality*, 40(1): 84-96. <http://doi.org/10.1016/j.jrp.2005.08.007>

- Chikara Hashimoto, Sadao Kurohashi, Daisuke Kawahara, Kenji Shinzato and Masaaki Nagata. 2011. Construction of a blog corpus with syntactic anaphoric and sentiment annotations. *Japanese Journal of Natural Language Processing*, 18(2): 175-201.
- Zhiheng Huang, Wei Xu, and Kai Yu. 2015. Bidirectional LSTM-CRF models for sequence tagging. arXiv pre-print arXiv:1508.01991.
- Ritsuko Iwai, Takatsune Kumada, Daisuke Kawahara, and Sadao Kurohashi. 2017. Translating Big Five personality constructs from English to Japanese, using statistical machine translation. Poster presented at the Eighteenth Annual Meeting of Society for Personality and Social Psychology.
- Ritsuko Iwai, Takatsune Kumada, Daisuke Kawahara, and Sadao Kurohashi. 2018. Automatic conceptual translation approach: Construction and validation of a Japanese Five-Factor personality questionnaire based on statistical machine translation. Poster presented at the Nineteenth Annual Meeting of Society for Personality and Social Psychology.
- Niklas Jakob and Iryna Gurevych. 2008. Extracting opinion targets in a single-and cross-domain setting with conditional random fields. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pages 1035-1045.
- Nobuhiro Kaji and Masaru Kitsuregawa. 2006. Automatic construction of polarity-tagged corpus from HTML documents. In *Proceedings of the 21st International Conference on Computational Linguistics*, pages 452-459.
- Kazuhito Kawabe, Makoto Miwa, and Yutaka Sasaki. 2015. Transportation terminology recognition for semi-automatic traffic ontology expansion. In *Proceedings of the Twenty-first Natural Language Processing in Japan*, pages 135-138.
- John Lafferty, Andrew McCallum, and Fernando Pereira. 2001. Conditional Random Fields: Probabilistic models for segmenting and labeling sequence data. In *Proceedings of Eighteenth Proceedings of the Eighteenth International Conference on Machine Learning*, pages 282-289.
- Robert R. McCrae, Alan B. Zonderman, Paul T. Costa, Michael H. Bond, and Sampo V. Paunonen. 1996. Evaluating replicability of factors in the Revised NEO Personality Inventory: Confirmatory factor analysis versus Procrustes rotation. *Journal of Personality and Social Psychology*, 70(3): 552-566. <http://doi.org/10.1037/0022-3514.70.3.552>
- Robert. R. McCrae, Antonio Terracciano, Members of the Personality Profiles of Cultures Project. 2005. Universal features of personality traits from the observer's perspective: Data from 50 cultures. *Journal of Personality and Social Psychology*, 88(3): 547-561. <http://doi.org/10.1037/0022-3514.88.3.547>
- Xuezhe Ma and Eduard Hovy. 2016. End-to-end sequence labeling via Bi-directional LSTM-CNNs-CRF. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, Volume 1: Long Papers*, pages 1064-1074.
- Naoto Nakazawa, Kanako Ikeda, Michika Yamada, Ryoma Yoshimura, Yui Suzuki, and Mamoru Komachi. 2018. Japanese phrase-unit corpus with polarity extracted from review documents. In *Proceedings of the 24th Annual Conference of Japanese Natural Language Processing*, pages 781-784.
- Tetsuya Nasukawa, Koichi Kamijo, Masahiro Yamamoto, and Hideya Kitamura. 2016. Investigation of linguistic features for personality estimation of author in Japanese. In *Proceedings of the Twenty-second Annual Conference of Japanese Natural Language Processing*, pages 1181-1141.
- Tetsuya Nasukawa and Koichi Kamijo. 2017. Estimating authors' personality in Japanese. In *Proceedings of the Twenty-third Annual Conference of Japanese Natural Language Processing*, pages 807-810.
- Shigeru Oishi, Jamie Schiller, and E. Blair Gross. 2012. Felt understanding and misunderstanding affect the perception of pain, slant, and distance. *Social Psychological and Personality Science*, 4(3): 259-266. <http://doi.org/10.1177/1948550612453469>
- Gregory Park, H. Andrew Schwartz, Johannes C. Eichstaedt, and Margaret L. Kern, Michal Kosinski, David J. Stillwell, Lyle H. Ungar and Martin E. P. Seligman. 2015. Automatic personality assessment through social media language. *Journal of Personality and Social Psychology*, 108(6): 934-952. <http://doi.org/10.1037/pspp0000020>
- Barbara Plank and Dirk Hovy. 2015. Personality traits on Twitter -or- how to get 1,500 personality tests in a Week. In *Proceedings of the 6th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, pages 92-98.
- Nils Reimers and Iryna Gurevych. 2017. Reporting score distributions makes a difference: Performance study of LSTM-networks for sequence tagging. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 338-348.

- William Revelle and David M. Condon. 2015. A model for personality at three levels. *Journal of Research in Personality*, 56: 70–81.
<http://doi.org/10.1016/j.jrp.2014.12.006>
- Gerard Saucier & Lewis R. Goldberg. 2001. Lexical studies of indigenous personality factors: Premises, products, and prospects. *Journal of Personality*, 69(6): 847–879. <http://doi.org/10.1111/1467-6494.696167>
- H. Andrew Schwartz, Johannes C. Eichstaedt, Margaret L. Kern, Lukasz Dziurzynski, Stephanie M. Ramones, Megha Agrawal, Achal Shah, Michal Kosinski, David Stillwell, Martin E. P. Seligman, and Lyle H. Ungar. 2013. Personality, gender, and age in the language of social media: The open-vocabulary approach. *PLoS ONE*, 8(9):e73791. <https://doi.org/10.1371/journal.pone.0073791>
- Pontus Stenetorp, Sampo Pyysalo, Goran Topić, Tomoko Ohta, Sophia Ananiadou, and Jun'ichi Tsujii. 2012. BRAT: A web-based tool for NLP-assisted text annotation. In *Proceedings of the Demonstrations at the Thirteenth Conference of the European Chapter of the Association for Computational Linguistics*, pages 102-107.
- Veselin Stoyanov and Claire Cardie. 2008. Topic Identification for Fine-Grained Opinion Analysis. In *Proceedings of the 22nd International Conference on Computational Linguistics*, pages 817-824
- Ryoji Suzuki, Makoto Miwa, and Yutaka Sasaki. 2015. SPARQL query conversion of question text for traffic ontology. In *Proceedings of the Natural Language Processing in Japan*, pages 171–174.
- Hirotoishi Taira, Takaaki Tanaka, and Masaaki Nagata. 2014. Construction of car license test corpus for textual entailment. In *Proceedings of the Twenty-eighth Annual Conference of the Japanese Society for Artificial Intelligence*, pages 1-3.
- Junya Takayama, Makoto Miwa, and Yutaka Sasaki. 2017. Constructing a Traffic Ontology-based QA Dataset. In *Proceedings of the Twenty-third Natural Language Processing in Japan*, pages 691-694.
- Tapus, A., Țăpuș, C., & Matarić, M. J. 2008. User—robot personality matching and assistive robot behavior adaptation for post-stroke rehabilitation therapy. *Intelligent Service Robotics*, 1(2): 169–183. <http://doi.org/10.1007/s11370-008-0017-4>
- Ryoko Tokuhisa, Kentaro Inui & Yuji Matsumoto. 2008. Emotion classification using massive examples extracted from the web. In *Proceedings of the 22nd International Conference on Computational Linguistics, Volume 1*, pages 881–888. <http://doi.org/10.3115/1599081.1599192>
- Ryoko Tokuhisa, Kentaro Inui & Yuji Matsumoto. 2009. Emotion classification: Using massive examples extracted from the web. *Japanese Journal of Information Processing* 50(4): 1365-1374.
- Toyota Nagoya Education Center Inc. 1998. *Driving Textbook (Rules and Knowledge)*.
- Toyota Nagoya Education Center Inc. 2004. *Driving Textbook (Driving Techniques)*.