Development of Methodologies and Language Resources for Acquiring Social Knowledge about Personality and Driving-related Behavior: The Synergy of Psychology and Natural Language Processing

Ritsuko Iwai

Abstract

In this dissertation, we developed methodologies and language resources for acquiring social knowledge on personality and driving-related behavior from large textual data. Social knowledge is expected to contribute to developing information systems, such as human–machine interaction (HMI) systems, which communicate naturally with humans. This acquisition entails the issues in both personality psychology and natural language processing (NLP). Despite newly emerging algorithms, personality psychology is yet struggling with integrating them into psychology and does not provide information that improves the implementation of HMI. Meanwhile, studies in NLP focus on algorithms and accuracy, but with restricted interest; they do not help shed light on the needs of human psychology. Synergizing psychology and NLP, we acquired social knowledge from large textual data collected from general populations in a manner that enables the knowledge to be applicable to informatics.

We designed the dissertation in a stepwise manner. Our approach required personality descriptors and behavioral expressions first. Throughout the dissertation, we carefully accumulated the integration of psychology and NLP.

Chapter 2 presents the related work in relation to Big Five personality, human behavior and Big Five personality, including driving behavior, and social knowledge and personality. Big Five is an internal framework in which personality is described in five broad traits: Extraversion (EX), Agreeableness (AG), Conscientiousness (CO), Neuroticism (NE), and Openness-to-Experiences (OP).

In Chapter 3, we proposed the *automatic conceptual translation approach* that adopts word embedding and statistical machine translation in the translation of English personality adjectives, and then developed and validated a new 20-item personality questionnaire (TDPI: *Trait Descriptors Personality Inventory*). The chapter has two purposes: to develop a Big Five questionnaire that retains conceptual equivalence between English Big Five adjectives and Japanese personality questionnaires, and to collect the Japanese personality descriptors for the dissertation. Our use of NLP techniques in questionnaire development is the first in both psychology and NLP, to the best of our knowledge. We presented results of six web surveys and one survey at one research institute (total respondents = 46,086). After selecting the final 20 items, we repeated the analyses to investigate reliability and validity. TDPI demonstrated consistently substantial internal consistency, substantial test-retest reliability (at two-week and 21-month intervals), factorial validity (repeated identifications of robust five-factor structure in different samples), criterion validity with other personality questionnaires, and

predictive validity. These results supported the construct validity of TDPI. Regarding the steps of development, we also acquired 116 unique personality descriptors to be used in the next chapter.

Chapter 4 describes the development of a personality dictionary based on human responses to TDPI and the candidate entries acquired with word embedding based on the personality descriptors acquired in Chapter 3. We automatically acquired 667 words and selected 526 based on human evaluations. We labeled the 317 words that directly describe individual personality as *personality trait words* (PTWs) and the 209 words that represent one aspect of personality but require additional words to describe personality as *personality-related words* (PRWs). We conducted a web survey (N = 1,938, female = 1,004, M_{age}= 49.8 years, SD = 16.3) with a one-week interval. The respondents completed TDPI and the 317 items of PTW (Time 1) and then another round of TDPI and the 209 items (Time 2). The PTW and PRW items were randomly presented. They responded to all of the items, including TDPI, using a seven-point Likert scale. We then examined TDPI's factorial validity and identified the similar five-factor patterns to the studies in Chapter 2. The results indicated substantial test-retest reliability (r = for EX, AG, CO, NE, and OP). We conducted exploratory factor analysis (EFA) individually, using TDPI+PTW (Time 1) and TDPI+PRW (Time 2). EFA of Time 1 indicated the similar five-factor pattern in which the TDPI items loaded on the hypothesized factors, whereas EFA at Time 2 indicated a different factor loading pattern. In calculating weights, we conducted multivariate single regression analyses for the entries of Time 1 based on the validity of the factor structure. For the entries of Time 2, we calculated Pearson's correlation coefficients.

In Chapter 5, we developed a *driving-related dictionary* (DRD) that includes psychological expressions, as a requirement for collecting textual data that describe human individual driving and the psychological experiences out of them in the following chapter. In the process of development, we first collected and listed 1,170 *driving-related words* (DRWs) and 836 *driving-behavioral words* (DBWs) from the following three resources: driving-related words developed by Nakagawa (2017), handwritten texts about driving (HWD) experiences and word entries in the domain of transportation in the dictionary of a Japanese morpheme analyzer, JUMAN. Using the DBWs, we automatically extracted predicate argument structures (PASs) from the following three resources: HWD, driving textbooks (Toyota Nagoya Education Center Inc., 1998; 2004) and Driving Behavior Corpus (Nakagawa, 2017). After filtering unrelated PASs, we asked human evaluators to label each PAS as follows: (A) direct behavioral expressions in driving, (B) contextual expressions in driving, (C) psychological expressions, and (D) situational expressions in driving.

Chapter 6 presents the construction of a driving experience corpus (DEC) (261 blog articles, 8,080 sentences) with four manually annotated tags, called driving experience (DE) tags, followed by three tags: others' behavior (OB), self-behavior (SB), and subjectivity (SJ). The dissertation focused on driving situations as they are part of people's daily lives and concern safety. A corpus annotated with behavior and subjectivity in driving situations is, thus, necessary. Subjectivity includes emotions, polarity, sentiments, human judgments, perceptions, and cognitions. We annotated spans with DE. Next, three tags, namely, OB, SB, and SJ, were annotated within DE spans. In addition to describing the guidelines, we presented corpus specifications, agreement between annotators, and three major difficulties during the development: the extended self, important information, and voice in mind. Automatic annotation experiments were conducted on the DEC using conditional random field (CRF) and bidirectional long short-term memory CRF (Bi-LSTM CRF). On the test set, the F-scores for both

CRF and Bi-LSTM CRF were about .55 for both OB and SB and approximately 75 for SJ. We provided error analysis of the CRF results, which revealed difficulties in interpreting nominatives and differentiating behavior from subjectivity.

Using all the resources developed in the previous chapters, we automatically extracted collocations of personality descriptors and behavioral PASs and then evaluated them as social knowledge by crowdsourcing. We trained the CRF model with DEC and applied the model to the Driving Behavior and Subjectivity Corpus (23,181 blog articles, 1.8 million sentences). We automatically extracted personality descriptors whose nominative are regarded as humans and extracted behavioral PASs under the condition that the PASs appear within the two sentences before and after the one with the personality descriptors. This extraction step yielded 5,345 collocations. Then, the collocations were evaluated through crowdsourcing to obtain social knowledge in four steps. First, we asked crowdworkers to judge whether a personality descriptor describes the human mind or psychological states. Second, crowdworkers filtered the analysis errors of predicates. Third, crowdworkers selected driving-related behaviors out of the filtered predicates. Finally, crowdworkers evaluated how “natural” the relations are between driving-related predicates and personality traits, to identify social knowledge.

Chapter 8 summarizes the findings, presents the contributions of the dissertation, discusses the synergy of psychology and NLP, and provides the limitations and the future directions for synergizing psychology and NLP. Finally, Chapter 9 concludes the dissertation.