2022-07-26

July 26, 2022 9:09 AM

Sentence Simplification with Deep Reinforcement Learning

- Recent approaches view the simplification process more holistically as a monolingual textto-text generation task borrowing ideas from statistical machine translation
- Simplification rewrites are learned automatically from examples of complex-simple sentences extracted from online resources such as the ordinary and simple English Wikipedia
- We experimentally show that the reinforcement learning framework is the key to successful generation of simplified text bringing significant improvements over strong simplification models across datasets.
- Neural Encoder-Decoder Model : The encoder-decoder model has two parts (see left hand side in Figure 1). The encoder transforms the source sentence X into a sequence of hidden states (h S 1 , h S 2 , . . . ,

h S |X|) with a Long Short-Term Memory Network (LSTM; Hochreiter and Schmidhuber 1997), while the decoder uses another LSTM to generate one word yt+1 at a time in the simplified target Y

Text Simplification for Reading Assistance: A Project Note

- 2.1 Readability assessment
  - a. Problem identification: identify which portions of a given text will be difficult for a given user to read,
  - b. Paraphrase generation: generate possible candidate paraphrases from the identified portions, and
  - c. Evaluation: re-assess the resultant texts to choose the one in which the problems have been resolved.
  - We have already proven that one can collect such readability assessment data by conducting survey questionnaires

targeting teachers at schools for the deaf.

- 2.2 Paraphrase acquisition
    - Manual collection of paraphrases in the context of language generation, e.g. (Robin and McKeown, 1996),
    - Derivation of paraphrases through existing lexical resources, e.g. (Kurohashi et al., 1999),
    - Corpus-based statistical methods inspired by the work on information extraction, e.g. (Jacquemin, 1999; Lin and Pantel, 2001), and
    - Alignment-based acquisition of paraphrases from comparable corpora, e.g. (Barzilay and McKeown, 2001; Shinyama et al., 2002; Barzilay and Lee, 2003).
    - One remaining

issue is how effectively these methods contribute to the generation of paraphrases in our application-oriented context.

- 2.3 Paraphrase representation
  - Our approach is (a) to define first a fully expressible formalism for representing paraphrases at the level of tree-to-tree transformation and (b) devise an additional layer of representation on its top that is designed to facilitate handcoding transformation rules.
  - 2.4 Post-transfer text revision
    - (1) The introduction of readability assessment would free us from cares about the

purposiveness of each paraphrasing rule in paraphrase acquisition.

- (2) Paraphrase acquisition is obviously indispensable for scaling up the resource.
- (3) A good formalism for representing paraphrasing rules would facilitate the manual refinement and maintenance of them.
- (4) Post-transfer error detection

and revision would make the system tolerant to flows in paraphrasing rules.

- While many researchers have addressed the issue of paraphrase acquisition reporting promising results as cited above, the other three issues have been left relatively unexplored in spite of their significance in the above sense

3 Readability assessment

- SLALOM is designed to capture the stereotypic linear order of acquisition within certain categories of morphological and/or syntactic features of language
- Unfortunately, the modeling method used in SLALOM cannot be directly applied to our domain for three reasons.
  - Unlike writing tutoring, in reading assistance, target sentences are in principle unlimited. We therefore need to take a wider range of morphosyntactic features into account.
  - SLALOM is not designed to capture the difficulty of any combination of morpho-syntactic features, which it is essential to take into account in reading assistance.
  - Given the need to consider feature

combinations, a simple linear order model that is assumed in SLALOM is unsuitable.

- 3.1 Our approach: We ask teachers
    - 3.1.1 Targets
        - We targeted teachers of Japanese or English literacy at schools for the deaf for the following reasons. Ideally, this sort of survey would be carried out by targeting the population segment in question, i.e., deaf students in our study

- In contrast, our approach is an attempt to model the knowledge of experts in this field (i.e., teaching deaf students). The targeted teachers have not only rich experiential knowledge about the language proficiency of their students but are also highly skilled in paraphrasing to help their students' comprehe nsion.

- Since such knowledge gleaned from individual experiences already has some generality, extracting it through a survey should be less costly and thus more comprehensive than investigation based on language proficiency testing.

- 3.1.2 Questionnaire
  - To make our questionnaire efficient for model acquisition

, we had to carefully control the variation in paraphrases. To do that, we first selected around 50 morpho-syntactic features that are considered influential in sentence readability for deaf people.

- 3.1.3 Administration
  - We administrated a preliminary survey targeting three teachers. Through the survey, we obser

ved that (a) the teach ers largel y agree d in their asses smen ts of relati ve reada bility, (b) their form at-free com ment s indic ated that the obser ved differ ences in reada bility were largel y expla inabl e in terms of the morp hosy ntacti c featu res we had prepa red, and (c) a larger scale d surve y was need ed to obtain a statistically reliable model.

- 3.2 Readability ranking model
  - The task of ranking a set of paraphrases can be decomposed into comparisons between two elements combinatorially selected from the set. We consider the problem of judging which of a given pair of paraphrase sentences is more readable/comprehensible for deaf students. More specifically, given paraphrase pair (si; sj), our problem is to classify it into either left (si is easier), right (sj is easier), or comparable (si and sj are comparable). Once the problem is formulated this way, we can use various existing

        techniques for classifier learning. So far, we have examined a method of using the support vector machine (SVM) classificatio n technique.
  - . 3.3 Evaluation and discussion
    - n To evaluate the two modeling methods, we conducted a ten-fold cross validation on the set of 4055 paraphrase pairs derived from the 770 questions used in the survey
    - The model achieved 95% precision with 89% recall
- 4 Paraphrase representatio
  - 4.1 Types of paraphrases of concern
    - There are various levels of paraphrase s as the following examples demonstrat e: (1) a. She burst into tears, and he tried to comfort her. b. She cried, and he tried to

console her.
    - (2) a. It was a Honda that John sold to Tom. b. John sold a Honda to Tom. c. Tom bought a Honda from John.
    - (3) a. They got married three years ago. b. They got married in 2000
  - Lexical vs. structural paraphrases Example (1) includes paraphrases of the single word “comfort” and the canned phrase “burst into tears”. The sentences in (2), on the other hand, exhibit structural and thus more general patterns of paraphrasing
  - . In developing a resource for paraphrasing, we have only to cover non-compositional (i.e., atomic) paraphrases. Compositional paraphrases can be handled if an additional computational mechanism for combining atomic paraphrases is devised
- Meaning-preserving vs. reference-preserving paraphrases
  - It is also useful to distinguish reference-preserving paraphrases from

meaningpreserving one

- 4.2 Dependency trees (MDSs)
    - Previous work on transfer-based machine translation (MT) suggests that the dependency-based representation has the advantage of facilitating syntactic transforming operations (Meyers et al., 1996; Lavoie et al., 2000). Following this, we adopt dependency trees as the internal representations of target texts. We suppose that a dependency tree consists of a set of nodes each of which corresponds to a lexeme or compound and a set of edges each of which represents the dependency relation between its ends
- 4.3 Three-layered representation
    - Previous work on transfer-based MT

systems (Lavoie et al., 2000; Dorna et al., 1998) and alignment-based transfer knowledge acquisition (Meyers et al., 1996; Richardson et al., 2001) have proven that transfer knowledge can be best represente d by declarative structure mapping (transformi ng) rules each of which typically consists of a pair of source and target partial structures as in the middle of Figure 2

- The SSR rule formalism allows a rule writer to edit rules with an ordinary text editor, which makes the task of rule editing much more efficient than providing her/him with a GUI-based complex tool for editing SR rules directly. The use of the extended natural language also has the advantage in improving the readability of rules for rule writers, which is particularly important in group work. To parse SSR rules, one can use the same parser as that used to parse input texts. This also improves the efficiency of rule developme nt because it significantly reduces the burden of maintaining the consistency between the POS-tag set used for parsing input and that used for rule specificatio ns.

- 5 Post-transfer error detection
  - The transfer errors observed in the experiment exhibited a wide range of variety from morphologi cal errors to semantic and

discourse-related ones.

- Most types of errors tended to occur regardless of the types of transfer. This suggests that if one creates an error detection module specialized for a particular error type, it works across different types of transfer. The most frequent error type involved inappropriate conjugation forms of verbs. It is, however, a matter of morphological generation and can be easily resolved.
- Errors in regard to verb valency and selectional restriction also tended to be frequent and fatal, and thus should have preference as a research topic.
- The next frequent error type was related

to the difference of meaning between near synonyms. However, this type of errors could often be detected by a model that could detect errors of verb valency and selectional restriction.

-