Automatic Question Generation Using Relative Pronouns and Adverbs

ABSTRACTS

This work presents a system that automatically generates multiple, natural language questions using relative pronouns and relative adverbs from complex English sentences. Our system is syntax-based, runs on dependency parse information of a single-sentence input, and achieves high accuracy in terms of syntactic correctness, semantic adequacy, fluency and uniqueness. One of the key advantages of our system, in comparison with other rule-based approaches, is that we nearly eliminate the chances of getting a wrong wh-word in the generated question by fetching the requisite wh-word from the input sentence itself. Depending upon the input, we generate both factoid and descriptive type questions. To the best of our information, the exploitation of wh-pronouns and wh-adverbs to generate questions is novel in the Automatic Question Generation task.
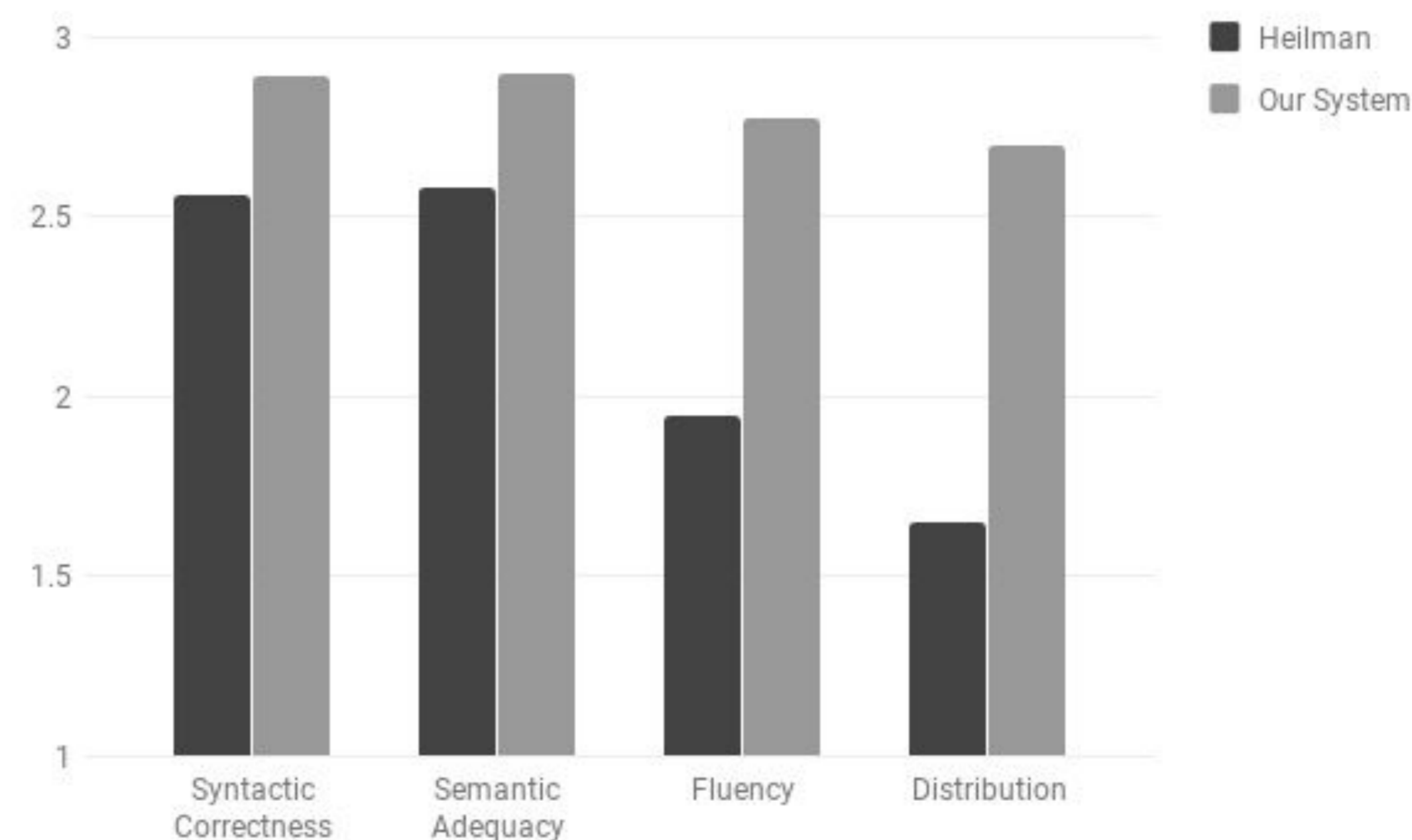
Sentences	Heilman	Our System
Mary gave dark chocolates and flowers to Bart who works at Amazon which was known to be the biggest online store in Africa	Who did Mary give dark chocolates and flowers to?	Who did Mary give dark chocolates and flowers to?
	Who gave dark chocolates and flowers to Bart who works at Amazon which was known to be the biggest online store in Africa?	Who works at Amazon?
	What did Mary give to Bart who works at Amazon which was known to be the biggest online store in Africa?	Who is Bart?
		What is the biggest online store in Africa?
	Did Mary give dark chocolates and flowers to Bart who works at Amazon which was known to be the biggest online store in Africa?	What is Amazon?

OBJECTIVE

In complex sentences, relative pronouns or relative adverbs perform the function of connecting or introducing the relative clause that is embedded inside the matrix clause. We aim to exactly exploit this structural relationship to generate questions, thereby adding to the pool of questions that can be generated from a given sentence. One of the key benefits of using the information from relative pronouns and relative adverbs is that we are not likely to go wrong with the wh-word, as we fetch it from the relative clause itself to generate the question.

METHOD

We split the complete QG task into the following sub parts - the input natural language sentence is first fed into the Spacy parser. Using the parse information, the system checks for the presence of one or more relative pronouns or adverbs in the sentence. Post that, it further checks for well defined linguistic features in the sentence, such as tense and aspect type of the root and relative clause verb, head-modifier relationship between different parts of the sentence, etc. to accordingly send the information to the rule sets. Depending upon which rule in the rule sets the information is sent to, questions are generated.

Authors: Konigari Rachna, Payal Khullar, Manish Srivastava Research Center: MT-NLP lab