Literacy situation models knowledge base creation

Uroš Škrjanc[1], Matthew Tanti[2], and Marko Novak[3]

Abstract

Keywords
story entities, relationship extraction

[1] *us1883@student.uni-lj.si, 63030323*
[2] *mt9734@student.uni-lj.si, 63210511*
[3] *mn3983@student.uni-lj.si, 63130166*

Introduction

The field of recognising the content and structure of texts and extracting content from them is increasingly related to machine learning methods. Not only syntax checking but also pure understanding of texts by humans is interesting for machine learning for different motives. There is practically no more field related to language that in some way could not be linked to machine learning methods.

Methods

In the group, we decided to try to analyse different short stories to extract information about the characters who appear in the texts and to find out what kind of relationships these characters are in.

First we found a collection of english short stories and choose a subset of them for training. Then, in the python programming language, we will use libraries such as NLTK, SpaCy, SNER, GATE... to determine which characters appear in the stories. At this stage, it will be necessary to extract the characters' actual names from the text and ensure that the correct numbers of characters and their positions in the text are extracted.

In the second part, we will try to make a model that properly connects characters in pairs into family relations. The model should determine whether a couple of characters are in a family relationship and, if so, then determine the type of relation. If the model finds that the persons are not in family relation, it will try to determine whether they are in any other relation (friends, business associate) or appear in the text independently of each other.

As part of the task, we will review the models that have already been implemented and, based on the results of the solutions already made, make our own model that would classify as accurately as possible.

Existing solutions

The use of deep neural networks is on the rise in already implemented solutions, but other approaches and combinations of these are also used in models for classifying family relations between persons in the text, such as rule-based approaches, utterance attribution and vocative detection technique and unsupervised approach to the extraction of interpersonal relations are described in the articles, which we intend to take as a basis for further work. [1] [2] [3] [4] [5] [6] [7] [8]

References

[1] Tommaso Caselli and Piek Vossen. The event storyline corpus: A new benchmark for causal and temporal relation extraction. In *Proceedings of the Events and Stories in the News Workshop*, pages 77–86, 2017.

[2] Shingo Nahatame. Revisiting second language readers' memory for narrative texts: the role of causal and semantic text relations. *Reading Psychology*, 41(8):753–777, 2020.

[3] Tom Trabasso and Paul Van Den Broek. Causal thinking and the representation of narrative events. *Journal of memory and language*, 24(5):612–630, 1985.

[4] Nasrin Mostafazadeh, Nathanael Chambers, Xiaodong He, Devi Parikh, Dhruv Batra, Lucy Vanderwende, Pushmeet Kohli, and James Allen. A corpus and cloze evaluation for deeper understanding of commonsense stories. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 839–849, 2016.

[5] Tirthankar Dasgupta, Rupsa Saha, Lipika Dey, and Abir Naskar. Automatic extraction of causal relations from text using linguistically informed deep neural networks. In *Proceedings of the 19th Annual SIGdial Meeting on Discourse and Dialogue*, pages 306–316, 2018.

[6] Danielle S McNamara and Arthur C Graesser. Coh-metrix: An automated tool for theoretical and applied natural language processing. In *Applied natural language processing: Identification, investigation and resolution*, pages 188–205. IGI Global, 2012.

[7] Rolf A Zwaan, Joseph P Magliano, and Arthur C Graesser. Dimensions of situation model construction in narrative comprehension. *Journal of experimental psychology: Learning, memory, and cognition*, 21(2):386, 1995.

[8] Rolf A Zwaan. Situation models: The mental leap into imagined worlds. *Current directions in psychological science*, 8(1):15–18, 1999.