先端科学技術研究科　修士論文要旨

所属研究室 (主指導教員)	ソーシャル・コンピューティング (荒牧　英治 (教授))		
学籍番号	2211147	提出日	令和 6年 1月 17日
学生氏名	清水　聖司		
論文題目	Clinical Temporal Relation Extraction: Accelerating Real-world Analysis by Efficient QA-based Annotations **Clinical Temporal Relation Extraction: Accelerating Real-world Analysis by Efficient QA-based Annotations**		

要旨

Temporal relation annotation in the clinical domain is crucial yet challenging due to its workload and the medical expertise required.

In this thesis, we propose a novel annotation scheme that integrates event start-points ordering and question-answering (QA) as the annotation format.

By focusing only on two points on a timeline, start-points ordering reduces ambiguity and simplifies the relation set to be considered during annotation.

QA as annotation recasts temporal relation annotation into a reading comprehension task, allowing annotators to use natural language instead of the formalisms commonly adopted in temporal relation annotation. Based on our scheme, most of the relations in a document are inferable from a significantly smaller number of explicitly annotated relations, showing the efficiency of our proposed scheme. Using these inferred relations, we develop a temporal relation classification model that achieves a 0.72 F1 score.

Also, by decomposing the annotation process into QA generation and QA validation, our scheme enables collaboration among medical experts and non-experts. We obtained high inter-annotator agreement (IAA) scores, which indicate the positive prospect of such collaboration in the annotation process. We plan to release the annotated corpus, annotation tool, and trained model to the public.

Finally, we present a proof of concept study to demonstrate how the trained model can automatically extract drug-disease temporal relations, discovering potential causes of adversarial drug effects.