Criminal Investigation with Augmented Ontology and Link Prediction

Takanori Ugai
Fujitsu Limited
Kawasaki, Japan
ugai@fujitsu.com

***Abstract*—This paper shows that our basic approach to the Knowledge Graph Reasoning Challenge is to create an ontology of motives and means. We also use link prediction based on knowledge graph embedding techniques to solve the case. The ontologies on motives and means are based on hand-crafted in our previous work and are supplemented by an automated ontology created by linking Wikidata. As a result of preliminary experiments, we present an overview of the ontology created at this stage and show how to predict characters with motives for incidents from the knowledge graph of the novel The Adventure of the Speckled Band.**

***Index Terms*—Knowledge Graph, Ontology, Wikidata, Link Prediction**

I. Introduction

The Knowledge Graph Reasoning Challenge [1] has been held in Japan since 2018. The Challenge aims to promote the technology of explainable AI using knowledge graphs. The challenge is to use a dataset of knowledge graphs representing eight Sherlock Holmes mystery novels to estimate the culprit with a reasonable explanation. 1st International Knowledge Graph Reasoning Challenge(IKGRC)[1] is an international version of the challenge, in which participants are tasked with solving the same task using a revised dataset.

The study aimed to find the murderer by inferring who had the motive, opportunity and method. The ontology extended using Wikidata was combined with the original knowledge graph to infer the murderer, motive, and weapon by link prediction using knowledge graph embedding. The following sections of this paper will be devoted to Section II. In Section II of this paper, we describe the structure of the knowledge graph provided by the knowledge graph inference challenge, Section III presents the basic policy we have adopted to solve the case. In addition, Section IV describes the Section IV outlines the ontology we have created as a preliminary experiment, based on links to wikidata[2], and shows the results of our predictions for the knowledge graph of the novel "The Adventure of the Speckled Band"(Speckled Band).

II. Structure of the knowledge graph provided

The schema of the knowledge graph provided by the organisers is based on the division of the set of content into scenes and the graphical representation of the content of each scene and the relationships between the scenes. Each scene is described by the following basic properties. The properties differ from the common <subject, predicate, object> format, where the scene ID is used as the subject and the information related to that scene is summarised [2].

- subject: The person or thing that is the subject of the description of the scene
- hasPredicate: A predicate describing the content of the scene
- hasProperty: The nature of the person or thing that is the subject
- Objects with scene details: whom, where, when, what, how etc.
- Relationships between scenes: then, if, because, etc.
- time: The absolute time the scene took place (xsd:DateTime)
- source: The original text of the scene (Literals in English and Japanese)

In the provided knowledge graphs, the types Person, Place, Situation and Object are given, but Object is not given a more specific type.

III. Our Approach

To deduce the culprit, various kinds of evidence (physical evidence and circumstances), such as the circumstances of the damage, the appearance of the victim and the suspect, and their relationships, are accumulated and compared with various knowledge related to the crime. Multiple analyses are conducted to narrow down the suspects and deduce the culprit.

The clues for the identification of the perpetrator include

1) the analysis of motive
2) the analysis of opportunity
3) the analysis of method
4) the analysis of behaviour and knowledge

In this study, we try to find the murderer by inferring the person with the motive and method. We create ontologies of motives and methods of crime from dictionaries, add them to the knowledge graphs and apply link prediction based on knowledge graph embedding to determine motives and methods. We manually made the final overall judgment based on motive, opportunity, and method.

[1]https://ikgrc.org/2023/application.htm
[2]https://wikidata.org/

<http://kgc.knowledge-graph.jp/data/SpeckledBand/anger> a fjs:negative_emotion .
fjs:negative _emotion rdfs:label “negative emotion"@en .
fjs:negative _emotion a rdfs:Class .
fjs:negative _emotion rdfs:subClassOf fjs:emotion .
fjs:emotion a rdfs:Class .
fjs:emotion rdfs:label "emotion"@en .

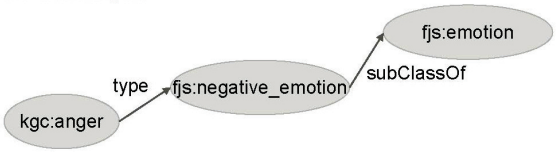

Fig. 1. Example of an object connected to Wikidatarepresenting

IV. Preliminary Experiment

A. Ontology Augmentation

In novels, Sherlock Holmes uses his knowledge and deductive powers, in addition to what he has seen and heard, to solve the case. The novel showcases these in the scene where the case is solved. Therefore, the knowledge graphs that represent the novel do not describe these knowledge and reasoning rules. Therefore, to explore the truth of the case formally and mechanically, it is necessary to add the knowledge that Sherlock Holmes would have about the crime's motives, means and alibis. In our previous work [3], we have created two ontologies that we use to investigate crimes. The first one is a structured ontology of criminal motives. Possible criminal motives were listed in the White Paper on Crime to create this ontology.

The means of killing were listed from the backward match of “kill” in the Japanese dictionary, and the details were described as structured data with attributes and values.

The previously created version was insufficient to cover the eight cases, so it was decided to extend it using Wikidata.

The extension of the knowledge graph with Wikidata was carried out in the following steps.

1) Extracting instances of object classes from the Novel Knowledge Graph.
2) Searching Wikidata for object strings.
3) Adding words close to each word as synonyms using Wikipedia word embedding data[3].
4) Assign relations and what he has seen and heard Wikidata's class hierarchy.

1069 classes were added, and 780 objects in the Novel Knowledge Graph were typed.

Figure1 is part of the ontology created[4]. From the knowledge graph of Speckled Band, extended by Wikidata, it can be deduced that Roma has negative feelings towards Roylott.

B. Case resolution using link prediction with knowledge graph embedding

In this section, we show the results of predicting who has motives for the knowledge graph of Speckled Band.

[3]https://fasttext.cc/docs/en/pretrained-vectors.html

[4]https://github.com/takanori-ugai/KGRCOntology

First, the triples obtained using the following SPARQL query are added to the original knowledge graph.

```
CONSTRUCT {
 ?who :hasMotivation ?motivation .
}
WHERE {
 ?scene :subject ?who .
 ?scene :hasPredicate ?motivation .
 ?motivation :instanceOf :Motivation .
}
```

Next, knowledge graph embedding is applied to the augmented knowledge graph. The simplest TransE [4] was used for the knowledge graph embedding. Finally, the motivation of each character is predicted and scored. For example, for Roylott, we predicted <Roylot hasMotivation ?motivation>, resulting in a score of 0.9 for property purpose and 0.8 for envy. Comparing each character's motivations, Roylott's property motive was plausible, followed by Helen's love motive.

V. Summary and Future Work

In this paper, we showed that our basic approach to the Knowledge Graph Reasoning Challenge is to create ontologies of motives and means. We also used link prediction based on knowledge graph embedding techniques to solve the case. The ontology on motives and methods is added to the previously hand-created ontology, automatically created by linking Wikidata. As a result of preliminary experiments, an overview of the ontology created at this stage is presented, showing how the knowledge graph of the novel “The Adventure of the Speckled Band” can be used to predict which characters have motives for the case and predict that Roylott is most likely to be a property seeker. This was shown to be the most likely prediction for Roylott. The same method will be applied to all novel knowledge graphs in the future, using the same way to predict how each character causes the incident.

References

[1] T. Kawamura, S. Egami, K. Matsushita, T. Ugai, K. Fukuda, and K. Kozaki, “Contextualized scene knowledge graphs for xai benchmarking,” *Proceedings of the 11th International Joint Conference on Knowledge Graphs (IJCKG2022)*, 11 2022.

[2] T. Kawamura, S. Egami, K. Tamura, Y. Hokazono, T. Ugai, Y. Koyanagi, F. Nishino, S. Okajima, K. Murakami, K. Takamatsu, A. Sugiura, S. Shiramatsu, X. Zhang, and K. Kozaki, “Report on the first knowledge graph reasoning challenge 2018,” in *Semantic Technology*, X. Wang, F. A. Lisi, G. Xiao, and E. Botoeva, Eds. Cham: Springer International Publishing, 2020, pp. 18–34.

[3] T. Ugai, Y. Koyanagi, and F. Nishino, “A logical approach to criminal case investigation,” *The 1st International Workshop on Knowledge Graph Reasoning for Explainable Artificial Intelligence*, 12 2021.

[4] A. Bordes, N. Usunier, A. Garcia-Duran, J. Weston, and O. Yakhnenko, “Translating embeddings for modeling multi-relational data,” in *Advances in Neural Information Processing Systems*, C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Weinberger, Eds., vol. 26. Curran Associates, Inc., 2013. [Online]. Available: https://proceedings.neurips.cc/paper/2013/file/1cecc7a77928ca8133fa24680a88d2f9-Paper.pdf