Criminal Investigation with Augmented Ontology and Link Prediction

22 January 2023

Takanori Ugai (Fujitsu Limited)

1. Information about the applicants

- Authors : Takanori Ugai

- Affiliations: Fujitsu Limited

- E-mail: ugai@fujitsu.com

2. Explanation of the reasoning and estimation process

■ Scope of knowledge used: All

■ Additional knowledge for reasoning and deduction

- Handmade motivational ontology
- Handmake means ontology

Extended and with

- Wordnet
- Cocept Net
- Word-embedding data created with fastText from Wikipedia

Performance information

Computer Specification

- CPU: 3.0GHz x16
- Memoty: 128GB
- Computation time:
  - 24 hours with link prediction and graph embedding.
  - 3 hours without link prediction

Reference Information

- T Ugai, Y Koyanagi, F Nishino: A Logical Approach to Criminal Case Investigation, The 1st International Workshop on Knowledge Graph Reasoning for Explainable Artificial Intelligence, December 2021, https://kgr4xai.ikgrc.org/papers/KGR4XAI_2021_paper_1_cr.pdf
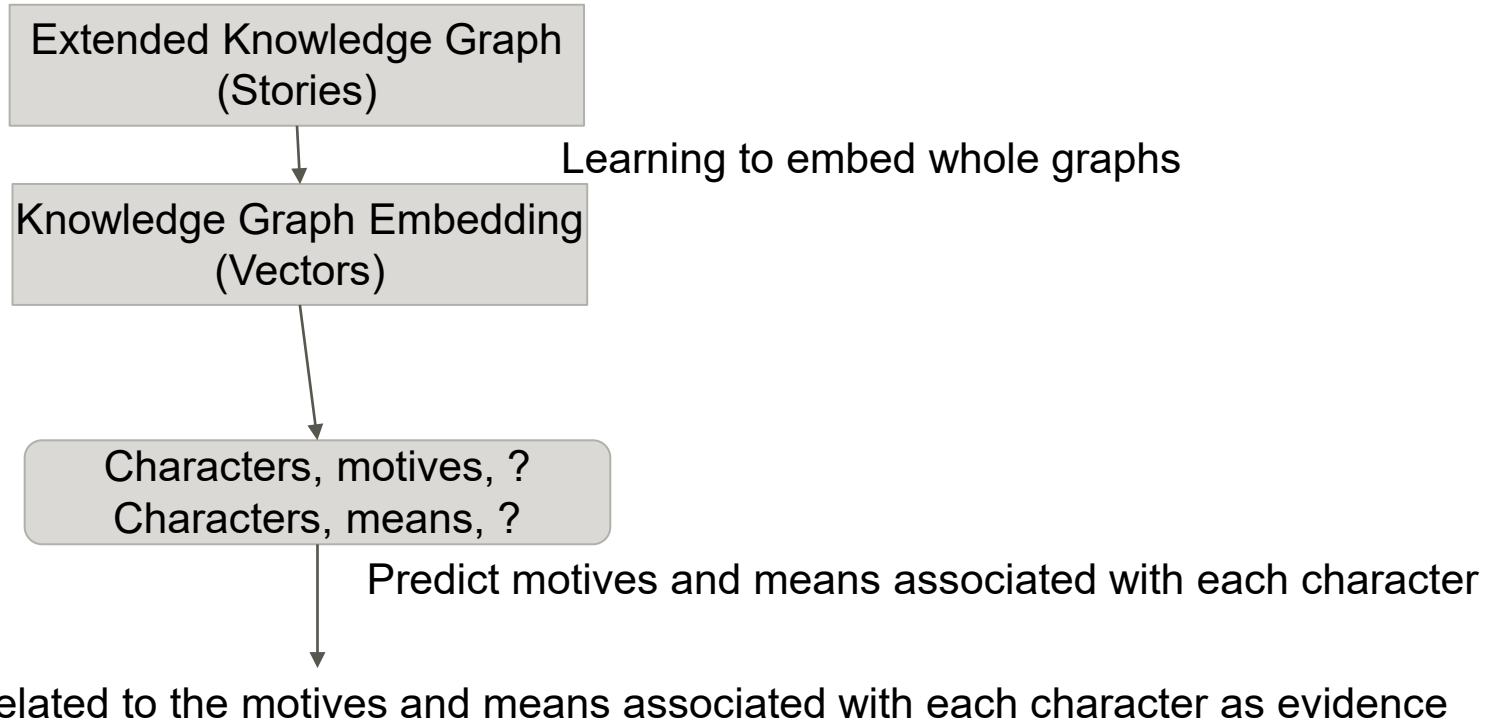
Basic strategy

- Inference of suspects, their motives, and means using ontology-based machine learning(knowledge graph embedding).

- Presenting scenes associated with motive and means as the basis for committing the offense.

Specific Description.

1. Linking motive ontology and means ontology with words appearing in the novel knowledge graph via wordnet (or Concept Net)

2. Predicting the most relevant motives and means for each character using graph embedding based on the enlarged knowledge graph.

3. Show scenes related to that character and the motives and means obtained.

Overall structure diagram (Augumented Knowlede Graph)

Motivation Dictionary
Envy, money,.

Extend with Wikipedia word embedding
and Wordnet graph embedding

Motive ontology
  Money: heritage, property
  Envy: righteous indignation, resentment

Linking

Wordnet, or Concept net

Linking

Augmented Knowledge Graph

Linking

Means ontology
 Poisoning: poison, viper
Stabbing: knife, javelin

Extend with Wikipedia word embedding
and Wordnet graph embedding

Murder Method Dictionary
Poisoning, stabbing

Learning and predicting culprits in the extended knowledge graph

Extended Knowledge Graph
(Stories)

Learning to embed whole graphs

Knowledge Graph Embedding
(Vectors)

Characters, motives, ?
Characters, means, ?

Predict motives and means associated with each character

Output scenes related to the motives and means associated with each character as evidence

Details

- Here we use The Adventure of the Speckled Band as an example to illustrate

- However, the ontology creation part is common throughout

Extension of the motive and method of murder dictionaries

- **Motivation Dictionary → Motivation Ontology**
  - envy instanceOf motive
  - money instanceOf motive
- **Murder-method dictionary -> Means ontology**
  - poisoning relatedWith snakes
  - stabbing relatedWith knife
  - shooting relatedWith gun

Structure of the Novel Knowledge Graph

- Scene 2.
  - Type Situation
  - Source 'Helen is scared'
  - Verb (state) scared
  - Subject Helen

  - rdf:type kgc:Situation;
  - kgc:source "Helen is scared"@en ;
  - kgc:source "Helen is scared"@en ;
  - kgc:hasPredicate
    <http://kgc.knowledgegraph.jp/data/SpeckledBand/beScared> ;
  - kgc:subject <http://kgc.knowledge-graph.jp/data/SpeckledBand/Helen>.

Extension of the Novel Knowledge Graph

- rdf:type kgc:Situation;
- kgc:source "Helen is scared"@en ;
- kgc:hasPredicate <http://kgc.knowledge-graph.jp/data/SpeckledBand/beScared>;
- kgc:subject <http://kgc.knowledge-graph.jp/data/SpeckledBand/Helen> ;
- :relatedWith "Helen"
- :relatedWith "is"
- :relatedWith "scared"

- Matching by string is inconvenient for further leveraging other knowledge and has poor search performance, so replace the string with an ID.
- In this case, we decided to use WordNet or Concept Net
  - In this material the case of Wordnet is described. Please refer the source codes for linking with concept net.
  - structure available, including links to meanings and synonyms
  - Easy to link to other dictionaries and resources such as wikidata

Link with Wordnet

- **<http://kgc.knowledge-graph.jp/data/SpeckledBand/2>**

  - rdf:type kgc:Situation;

  - kgc:source "Helen is scared"@en ;

  - kgc:hasPredicate <http://kgc.knowledge-graph.jp/data/SpeckledBand/beScared> ;

  - kgc:subject <http://kgc.knowledge-graph.jp/data/SpeckledBand/Helen> ;

  - :relatedWith <http://www.w3.org/2006/03/wn/wn31/instances/word-helen> ;

  - :relatedWith <http://www.w3.org/2006/03/wn/wn31/instances/word-be>;

  - :relatedWith <http://www.w3.org/2006/03/wn/wn31/instances/word-scared> .

Finding people to motivate

- Select ?subject, ?word, ?motivation where {
  - ?s kgc:subject ?subject .
  - ?s :relatedWith ?word .
  - ?word :relatedWith ?motivation .
  - ?motivation :instanceOf :Motivation .
- }

- Subject : Helen
- Word : scared
- Motivation: defence, self-defence

- Helen hasMotivation defence
- Helen hasMotivation self-defence

Predicting motive for each suspect.

- **Helen hasMotivation ?motivation .**
  - Resentment (0.4)
  - Money(0.3)
  - :

- **Roylott hasMotivation ?motivation .**
  - Money(0.8)
  - Love (0.6)
  - Resentment(0.3)

Search for scenes that lead to a motive.

- Select ?s ?word where {
-     ?s kgc:subject Roylott.
-     ?s :relatedWith ?word .
-     ?word :relatedWith Love.
- }
- kd:113 kgc:source " Roylott is not opposed to the marriage between Helen and Persie Armitage."@en ;
- kgc:relatedWith wn31instances:word-marriage

- The key word 'marriage' indicates a motive to kill out of love.

Predicting means for each suspect.

- Roylott hasMean ?mean .

- Poisonig(0.6)

- Strangling(0.3)

- :

Search for scenes connected to means.

- Select ?s ?word where {
-   ?s ?p Roylott.
-   ?s :relatedWith ?word .
- ?word :relatedWith Poisoning.
- }
- kd:84 kgc:source "On death day of Julia, Roylott gave Julia a drink of brandy"@en ;
- kgcf:relatedWord wn31instances:word-a , wn31instances:word-drink , wn31instances:word-give , wn31instances:word-brandy , wn31instances:word-on , wn31instances:word-death , wn31instances:word-day .

- Output of statements suggesting poisoning by drink.

Conclusion of the case:

- **The following possibilities were output**
  - Roylott killed for property purposes.
    - Means could include poisoning (poison in brandy), strangulation (bedroom string)
  - Helen killed for property
    - Means could include strangulation (bedroom strings)

- **In the future, mechanisms need to be built in to enable the appropriate determination of the object of motive.**

- **The Speckled Band:** Who killed Julia? (criminal & explanation)
- A Case Of Identity
- **The Devil's Foot:** Who killed the victims? (criminal & explanation)
- **The Crooked Man:** Why did Barclay die? (explanation)
- **The Dancing Men:** Break the codes (code breaking)
- **The Abbey Grange:** Who killed Lord Blackenstall? (criminal & explanation)
- **The Resident Patient:** Who killed Blessington? (criminal & explanation)
- **Silver Blaze:** Who took out the White Silver Blaze? (criminal & explanation)

Other 7 stories tasks

■ **The Speckled Band:** Who killed Julia? (criminal & explanation)

■ A Case Of Identity

■ **The Devil's Foot:** Who killed the victims? (criminal & explanation)

■ **The Abbey Grange:** Who killed Lord Blackenstall? (criminal & explanation)

■ **The Resident Patient:** Who killed Blessington? (criminal & explanation)

■ **Silver Blaze:** Who took out the White Silver Blaze? (criminal & explanation)

■ We applied our method to the 5 stories.

A Case Of Identity

- **The following possibilities were output**
  - policeman kidnapped out of love
  - Parents kidnapped for property
  - means were not found in particular

- **It is considered we can say that the parents were kidnapped for property**

The Devil's Foot

■ For this novel, neither the suspect nor any possible way of finding out about the suspect could be found.

The Abbey Grange

- **The following possibilities were output**
  - Policeman killed out of love
  - Soldier killed because of pleasure
  - Criminals killed because of resentment
  - Means output about the offender was that he was poisoned

- **have not been able to find a correct answer to this case**

The Resident Patient

- **The following possibilities were output**
  - Blessington killed out of money or love
    - Suggested method was burning.
  - Percy Trevelyan killed out of love or money
    - Suggested methods were stabbing or strangulation
  - Russian Aristocracy killed out of love
    - Suggested method was the eradication

- **From the situation, we can say the criminal is detected correctly**

Silver Blaze

- **The following possibilities were output**
  - Fitzroy_Simpson did this because of belief or drug
    - Suggested methods were sting or shooting, burning
  - Silas_Brown did this because of greed
    - Suggested methods were sting or shooting, burning

- **In the story, nobody was involved.**
- **Expected person is Silver Blaze, but Silver Blaze was not suggested.**

Summary

- **The Speckled Band:** the correct answer was suggested

- A Case Of Identity: a kind of correct answer was suggested.

- **The Devil's Foot:** no answer was suggested

- **The Abbey Grange:** a kind of correct answer was suggested

- **The Resident Patient:** the correct answer was not suggested

- **Silver Blaze:** the correct answer was not suggested

Discussion

■ For motive, we believe that the reason for the output of envy, money and love in all cases is that, when creating the motive ontology, these three were associated with words that are also used outside the context of the case.

■ For means and methods, we believe that this is because, when assigning a related word for xxx killings, the word xxx killings was associated in the same way rather than means (e.g. the related word for stabbing was gunshot).

■ when creating the respective ontology

  ■ to eliminate as far as possible the duplication of related terms in developing the means ontology

  ■ to associate words more closely with the means

■ are necessary.

3. Developed application

■ Download

  ■ https://github.com/takanori-ugai/Challenge2023

■ How to execute

  ■ praparation

    • git clone https://github.com/takanori-ugai/Challenge2023.git
    • cd Challenge2023
    • ./gradlew build
    • cd dict ; sh prepare.sh ; cd ..
    • cd data ; sh prepare.sh ; cd ..
    • set JAVA_OPTS="-Xms1g -Xmx16g"
    • make

  ■ execution

    • make SpeckledBand # in the case of SpeckledBand

Environments

- OS: CentOS 7.7

- JDK: openjdk version "18.0.2" 2022-07-19

- Knowledge Graph Embedding: Owned implemented TransE

4. Sharing of materials

- **[Application Sheet]**
  - Public availability:
    - ( ) Allow to publish
  - Publication format
    - ( ) Publish on your own site and wish to be linked to our site.
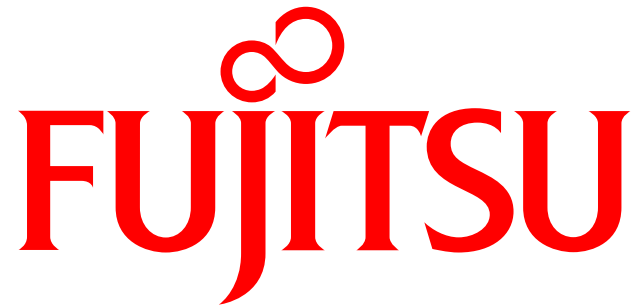   - -> URL*: I will tell you later
- **[Submitted application, source code, data, etc.]**
  - Public availability:
    - ( ) Allow to publish
  - Publication format
    - ( ) Publish on your own site and wish to be linked to our site.
      - -> URL*: https://github.com/takanori-ugai/Challenge2023/