

第2回ナレッジグラフ推論チャレンジ2019 応募シート

2019年12月3日

チーム名: 上小田中411

メンバー: 鵜飼孝典、岡嶋 成司

所属: 富士通研究所

代表のメールアドレス: ugai@fujitsu.com

1. 応募者に関する情報

- チーム名: 上小田中411
- 所属: 富士通研究所
- メールアドレス: ugai@fujitsu.com

2. 推論、推理過程の説明

- 利用した知識の範囲:すべて
- 推論、推理のために追加した知識
 - 動機オントロジ
 - 手段オントロジ前回応募時に作成したものを拡張して利用した
 - RDF版Wordnet
 - WikipediaからfastTextで作成した単語埋め込みデータ
- パフォーマンス情報
 - 計算機のスペック
 - CPU: 3.0GHz x16
 - メモリ: 128GB
 - 計算時間
 - 最初からすべて学習させるところから始めると18時間
 - 最後の犯人を推定する部分だけであれば、数分
- 参考情報
 - 昨年の応募作品
 - 上小田中411 <https://github.com/KGChallenge/Challenge>
 - FLL-ML <https://github.com/okajima-s/kgc2018-fllml/>

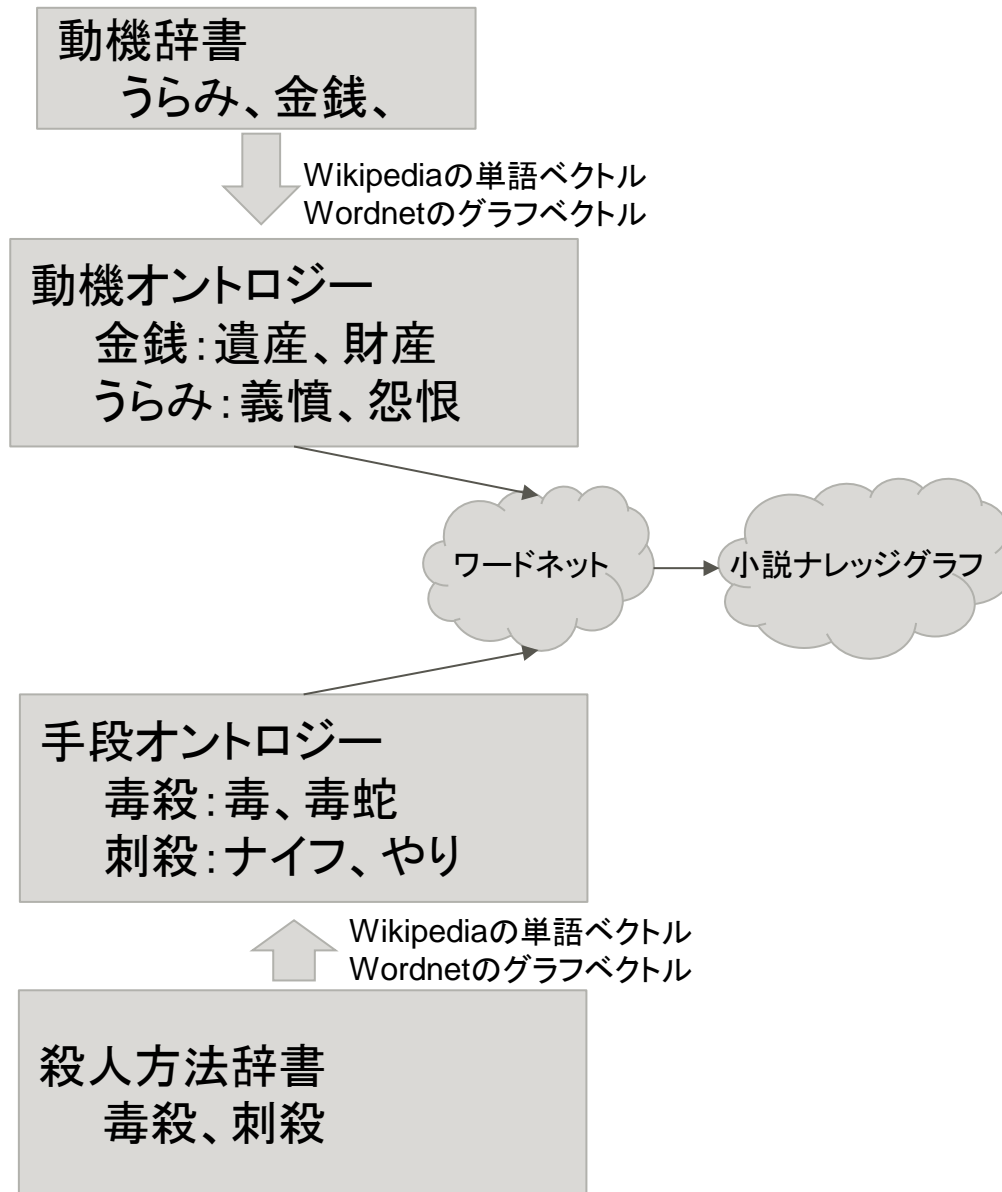
■ 基本戦略:

- 容疑者とその動機、手段をオントロジーをもちいた機械学習で推論
- その動機、手段に関連付けられたシーンをその容疑者が犯罪を犯した根拠として示す

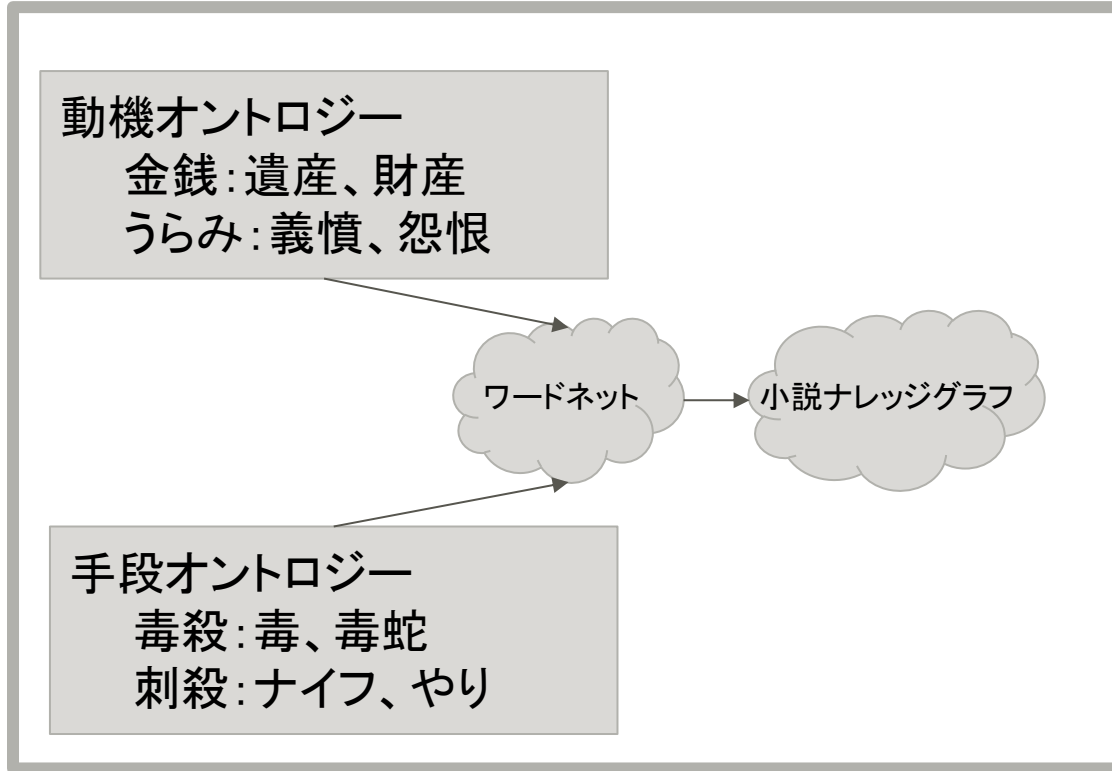
■ 具体的説明

1. 動機オントロジーと手段オントロジーを小説ナレッジグラフに出現する単語をwordnetを経由してリンク付けする
2. 拡大したナレッジグラフをもとにグラフ埋め込みを用いて、各登場人物にもっとも関連する動機と手段を予測する
3. その登場人物と得られた動機、手段に関連するシーンを示す

全体構成図(拡大ナレッジグラフ)



拡大ナレッジグラフの学習と犯人予測



全体のグラフ埋め込みを学習

登場人物、動機、?
登場人物、手段、?

各登場人物に関連する動機、手段を予測



各登場人物に関連する動機、手段に関連するシーンを根拠として出力

具体的手順

- ここでは、まだらのひもを例に用いて説明する
- ただし、オントロジー作成の部分は、全体で共通

■ 動機辞書→動機オントロジー

- うらみ instanceof 動機
- 金銭 instannceOf 動機
- お金 relatedWith 金銭

■ 殺人方法辞書→殺人方法オントロジー

- 毒殺 relatedWith 蛇
- 刺殺 relatedWith ナイフ
- 射殺 relatedWith 銃

- シーン2
 - タイプ シチュエーション
 - ソース 「ヘレンは怖がっている」
 - ソース 「Helen is scared」
 - 動詞(状態) 恐れている
 - 主語 ヘレン
-
- <http://kgc.knowledge-graph.jp/data/SpeckledBand/2>
 - rdf:type kgc:Situation;
 - kgc:source "ヘレンは怖がっている"@ja ;
 - 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>.

- <http://kgc.knowledge-graph.jp/data/SpeckledBand/2>
- rdf:type kgc:Situation;
- kgc:source "ヘレンは怖がっている"@ja ;
- 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"

- 文字列で一致させるのは、さらに他の知識を活用するのに不便ですし、検索性能も悪いので、文字列をIDに置き換えます。
- 今回は、WordNetを利用することにしました
 - 意味や同義語へのリンクなどの構造を利用可能
 - 他の辞書やwikidataなどのリソースとのリンクも容易

- <<http://kgc.knowledge-graph.jp/data/SpeckledBand/2>>
- rdf:type kgc:Situation;
- kgc:source "ヘレンは怖がっている"@ja ;
- 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>> .

動機につながる人を探す

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

- Subject : ヘレン
- Word : scared (恐れ)
- Motivation: 防衛、護身

- Helen hasMotivation 防衛
- Helen hasMotivation 護身

- などのここで得られた登場人物と動機データを追加しました

■ ヘレン hasMotivation ?motivation .

- うらみ (0.4)
- 金銭(0.3)
- :

■ ロイロット hasMotivation ?motivation .

- 愛(0.6)
- うらみ(0.3)
- :

動機につながるシーンを検索する

- Select ?s ?word where {
- ?s kgc:subject ロイロット.
- ?s :relatedWith ?word .
- ?word :relatedWith 愛.
- }

- kd:113 kgc:source "ロイロット博士はヘレンとパーシィ・アーミ
 ティッジの結婚に反対しない。"@ja , " Royslott is not opposed to
 the marriage between Helen and Persie Armitage."@en ;
 kgc:relatedWith wn31instances:word-marriage

- 結婚というキーワードで、愛ゆえに殺すという動機を示している

■ ロイロット hasMean ?mean .

- 毒殺(0.6)

- 絞殺(0.3)

- :

手段につながるシーンを検索する

- Select ?s ?word where {
- ?s ?p ロイロット.
- ?s :relatedWith ?word .
- ?word :relatedWith 毒殺.
- }

- kd:84 kgc:source "事件当夜, ロイロットはジュリアにブランデーを飲ませた"@ja , "On death day of Julia, Royslott 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 .

- 飲み物で毒殺を示唆する文を出力した

■ 以下のような可能性が出力された

■ ロイロットが財産目的で殺した

- ・ 手段としては、毒殺（ブランデーに毒）、絞殺（寝室のひも）などが考えられる

■ ヘレンが財産目当てで殺した

- ・ 手段としては、絞殺（寝室のひも）などが考えられる

■ 今後、動機の対象を適切に判断できるような仕組みを組み込む必要がある

「まだらのひも」以外の小説の取り扱い

- 以下のように5つの事件を犯罪として、犯人と動機、手段を予測し、関連するシーンを出力する: 詳細は、githubを参照してほしい

- まだらのひも : ジュリアの死因を探る
 - ・ 誰がどんな理由で殺したのか？
 - ・ どうやって殺したのか？
- 踊る人形: キュービット氏と妻の死因を探る
 - ・ 誰がどんな理由で殺したのか？
 - ・ どうやって殺したのか？
- 花婿失踪事件: ホズマーの失踪の原因を探る
 - ・ 誰がどんな理由で失踪させたのか？
 - ・ どうやって失踪させたのか？
- 悪魔の足: ブレンダとモーティマーの死因を探る
 - ・ 誰がどんな理由で殺したのか？
 - ・ どうやって殺したのか？
- 背中曲がった男: ジェームズの死因を探る
 - ・ 誰がどんな理由で殺したのか？
 - ・ どうやって殺したのか？

- 以下のような可能性が出力された
 - 警官が愛ゆえにさらった
 - 親が財産目当てでさらった

- 手段は、特に見つからなかった

- 親が財産目当てでさらったというのは、あてたと言ってもよいと考えられる

事件の結論：背中の曲がった男

- 以下のような可能性が出力された
 - 警官が愛ゆえに殺した
 - 兵士が喜びゆえに殺した
 - 犯罪者が、恨みゆえに殺した
- 手段は、犯罪者について、毒殺したということが出力された
- この事件については、正解を見つけることができていない

- この小説については、容疑者もその容疑者についての可能性のある方法も見つけることができなかった

- この小説については、容疑者もその容疑者についての可能性のある方法も見つけることができなかった

- まだらのひもについては、
 - 容疑者としては、ロイロット、ヘレン
 - 手段としては、ロイロットの毒殺、ロイロット、ヘレンの寝室にあるひも(ロープ)による絞殺
- が候補として上位に出力された
- 花嫁失踪事件については、親が誘拐犯人として推論された。
- 他の3つについては、容疑者も手段も決定的なものは出力されなかった。

- 動機についてどの事件も、うらみ、金銭、愛が出力される原因としては、動機オントロジー作成の際、これらの3つについては、事件の文脈以外でも使われる単語が関連付けられたためと考えている
- 手段、方法については、〇〇殺について、関連語を付与する際、手段よりも、同じように××殺という語が関連付けられたためと考えている（例えば、刺殺の関連語に銃殺が含まれた）
- それぞれのオントロジーを作成する際
 - 関連語の重複をできるだけなくすること
 - 手段オントロジーでは、より手段に近い語を関連付けられるような展開をおこなうこと
- が必要であると考えている

3. 実行プログラム

■ ダウンロードできる場所

- <https://github.com/KGChallenge/Challenge2019>

■ 実行方法の説明

- `git clone https://github.com/KGChallenge/Challenge2019.git`
- `cd Challenge2019`
- https://drive.google.com/open?id=1UN8_bUo4qvWUyb-tUS6aq_xP0_3pmxmW から `dict.tar.gz` と `data.tar.gz` をダウンロード
- `tar xf dict.tar.gz`
- `tar xf data.tar.gz`
- `make SpeckledBand` #「まだらのひも」の場合
 - ・「まだらのひも」以外の小説については、github のサイトをご参照ください

■ 既存ツールの詳細(動作確認環境)

- OS: CentOS 7.6
- JDK: Java(TM) SE Runtime Environment (build 1.8.0_221-b11)
- Jena : apache-jena-3.13.0
- グラフ埋め込みはJavaで独自に実装したTransEを用いている

4. 資料の共有について

■ 応募フォーム

■ 公開してよい

■ 公開形式

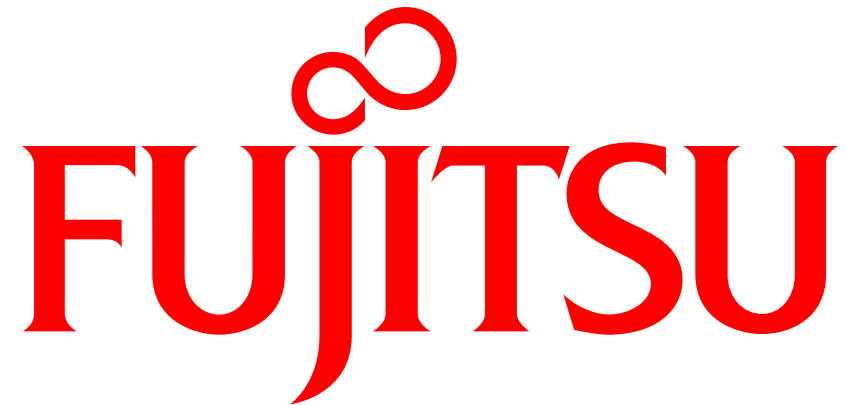
- 独自のサイトで公開してリンクを希望
 - のちほど slideshare の URL をお知らせします

■ 応募フォーム

■ 公開してよい

■ 公開形式

- 独自のサイトで公開してリンクを希望
 - <https://github.com/KGChallenge/Challenge2019>



shaping tomorrow with you