

第1回ナレッジグラフ推論チャレンジ2018開催報告

～説明性のある人工知能システムを目指して～

Report on the First Knowledge Graph Reasoning Challenge 2018

- Toward the eXplainable AI System -

川村 隆浩 *1*2 Takahiro Kawamura	江上 周作 *2 Shusaku Egami	田村 光太郎 *3*4 Koutarou Tamura	外園 康智 *4 Yasunori Hokazono	鷗飼 孝典 *5 Takanori Ugai
小柳 佑介 *5 Yusuke Koyanagi	西野 文人 *5 Fumihito Nishino	岡嶋 成司 *5 Seiji Okajima	村上 勝彦 *5 Katsuhiko Murakami	高松 邦彦 *6 Kunihiko Takamatsu
杉浦 あおい *7 Aoi Sugiura	白松 俊 *8 Shun Shiramatsu	張 翔宇 *8 Shawn Zhang	古崎 晃司 *9 Kouji Kozaki	

*1 科学技術振興機構 Japan Science and Technology Agency	*2 電気通信大学 University of Electro-Communications	*3 NRI デジタル NRI digital, Ltd.
*4 野村総合研究所 Nomura Research Institute, Ltd.	*5 富士通研究所 Fujitsu Laboratories Ltd.	*6 神戸常盤大学 Kobe Tokiwa University
*7 神戸市立西神戸医療センター Kobe City Nishi-Kobe Medical Center	*8 名古屋工業大学 Nagoya Institute of Technology	*9 大阪大学 Osaka University

New challenges for knowledge graph reasoning started in Japan in 2018. Machine learning techniques have promoted applying systems containing AI to a wide variety of social systems. The explanation for the AI decision is becoming important to guarantee the security and safety of such AI systems to a certain extent. Therefore, the Special Interest Group on Semantic Web and Ontology of JSAI organized this challenge, which calls for techniques for reasoning and/or estimating criminals with a reasonable explanation based on an open knowledge graph representing a well-known novel of Sherlock Holmes. This paper describes a summary report of the first challenge, including the knowledge graph constructed for this task, approaches proposed for reasoning and/or estimating, and the evaluation metrics. Finally, we conclude this paper with the plans and issues for the next challenge in 2019.

1. チャレンジ開催の背景と目的

近い将来、さまざまな社会システムに AI 技術が組み込まれるようになることが期待される。いずれは人の手を離れてシステム自身に決定的な判断を任せられるようになるだろう。しかし、安全・安心に社会の中で AI を活用していくためには、AI によるシステムが正しく動作しているかを確認する必要がある。だが、深層学習などの機械学習技術は判断や予測のプロセスがブラックボックスであり、結論に至った理由を人間は理解できていない。そのため、AI システムが結論に至った理由を説明できる性質、説明性 (explainability) または解釈性 (interpretability) を有する AI 技術が注目されている。また、現在の AI 技術は機械学習のアプローチが主であるが、今後は帰納的な機械学習と演繹的な知識活用を融合した形の AI 技術が必要になると考えられる。しかし現状、帰納的な機械学習技術と演繹的な知識活用技術を適切に評価するためのデータセットは存在しない。推論・推定技術を用いて解釈可能な AI を実現するためのタスクに適したデータセットとしては、現実社会を反映したより複雑な、例えば時間的、因果関係的、確率的関係性を含んでいることが必要である。そこで、本会セマンティック Web とオントロジー (SWO) 研究会では、説明性、解釈性を有する AI 技術、特に帰納的な機械学習 (推定) と演繹的な知識活用 (推論) を融合した AI 技術の開発に向けたチャレンジの開催を決定した [川村 18]。本チャレンジは、(1) 推

論・推定技術共通の評価セットとして、現実の社会問題や人間関係など複雑で構造的な関係性を含んだ大規模ナレッジグラフを構築し、(2) オープンサイエンスの一環としてそれらを公開した上で、幅広く IT 技術者から推論・推定に関する手法を結集し、説明性、解釈性に関する適切な指標を設計した上で、客観的な評価、分類、体系化を行うことを目的としている。

なお、本チャレンジで構築したナレッジグラフ、応募作品、その他関連資料は公式サイト *1 に掲載されている。

2. ナレッジグラフの構築

1 回目となった 2018 年度は、シャーロック・ホームズの短編推理小説 *2 「まだらのひも」を題材とし、事件や背景、人物像を知識化したナレッジグラフで与えられる情報に基づいて事件の犯人を正しく突き止め、その理由 (証拠やトリックなど) を適切に説明するタスクとした。推理小説とした理由は、

- 現実社会の複雑な関係性を含みながら仮想的にクローズな (答えがあり、それに至る制約を制御できる) タスクを設計できる、
- タスクによっては不確実情報や証拠写真など確率的な処理や機械学習を入れないと解けなかったり、陽に書かれていない常識知識を補完しなくては解けない等、推定と推論の融合を促せる、

連絡先: 川村 隆浩, 国立研究開発法人科学技術振興機構, takahiro.kawamura(at)jst.go.jp

*1 <http://challenge.knowledge-graph.jp>

*2 <https://ja.wikipedia.org/wiki/Category:シャーロック・ホームズシリーズの短編小説>

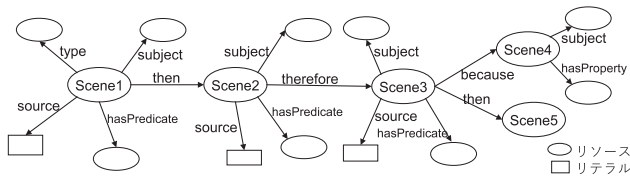


図 1: ナレツジグラフのイメージ

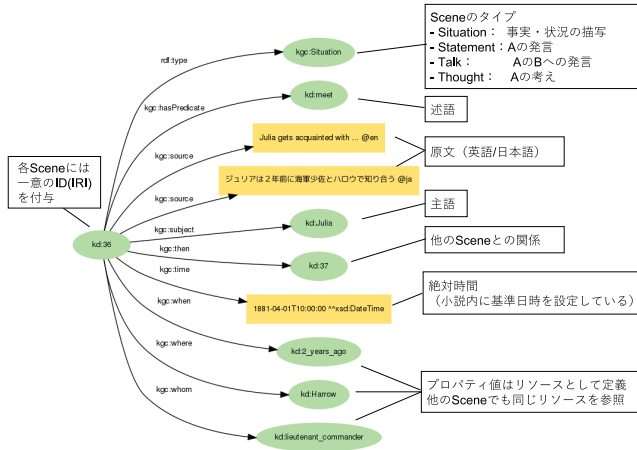


図 2: シーンの記述例

- 読者が納得しないと小説として成立しないという人間に対する説明性を有している,
- 小説が広く一般に知られており関心を引きやすい,

ことなどが挙げられる。

2.1 ナレツジグラフスキーマの詳細

ナレツジグラフのスキーマとしては一連の内容を最小単位の場面(シーン)に分割し、各シーンの記述内容、およびシーン間の関係をグラフ化する形とした。図 1 にナレツジグラフの全体イメージを示す。

また、各シーンの記述に用いる基本的なプロパティとして以下を用意した。シーンに付随する情報をまとめるため、シーン ID を主語としたプロパティとなっており、一般的な <subject, predicate, object> の形式ではないことに注意が必要である。図 2 にシーンの記述例を示す。

- subject: その場面の記述において主語となる人やもの
- hasPredicate: その場面の内容を表す述語
- hasProperty: 主語となる人やもののもつ性質
- 場面の詳細を表す目的語: whom (誰に), where (どこで), when (いつ), what (何を), how (どのように) など
- 場面間の関係: then, if, because など
- time: その場面が起こった絶対時間 (xsd:DateTime)
- source: その場面の原文 (英語/日本語のリテラル)

ナレツジグラフは、研究データのオープン化に向けて世界的に広く用いられている Resource Description Framework (RDF) で記述されており、グラフ DB に対する標準的なクエリ言語 SPARQL にて検索可能である。

3. 推論・推定技術への取り組み

この問題に対して独自性に富んだ 8 件の取り組み(実装を伴う応募が 5 件, アイデアのみの応募が 3 件)が提案された。以下、受賞作品を中心に 4 つの取り組みを簡単に紹介する。各

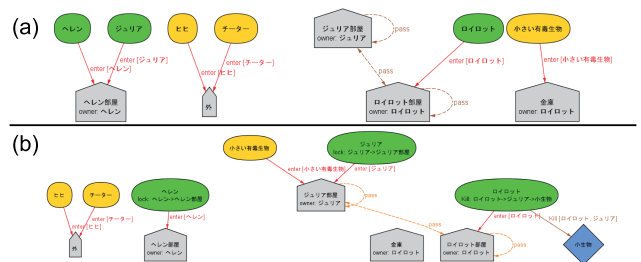


図 3: ジュリア殺害の日における整合解のうちの一つ。犯行当日の (a) 夜と (b) 深夜。小さい有毒生物が夜の段階でロイロット部屋内にいるのか、金庫内にいるのかは情報がなく特定できないため、解が一意に定まらない。

作品の詳細については、前述の公式サイトの結果発表ページを参照してほしい。

3.1 野村総合研究所の取り組み

犯行状況を数理的に記述するために、本文だけでは欠損している知識の補完を、関係性データの解析で行い、成立可能な状況の探索は SAT 問題に落とし込むことを考えた。

本手法で大きく貢献した方法については次の通りである。詳細なスクリプトや条件設定は、公式サイトの結果発表ページに掲載の資料を参照してほしい。

まず、テンソル解析による知識補完では、主語・述語・目的語の SVO 形式 (SVOO は 2 つの SVO に分解) の 3 軸とし、Tucker 分解によって補完した。補完された知識には、ジュリアに対して Animal が何らかの行為を行ったとの知識が出現した。数少ないジュリアを対象とする行動の情報として有力と判断した。

次に、所与の文章を述語論理式に書き直した。場面に共通する登場主体、殺人方法、建物/所在に関する情報を述語論理式として表し、場面「ジュリア殺害の日」、「ヘレン殺害未遂の日」における人物や状況を事実として、犯行状況の可能性を SAT 問題の解として列挙を行った (図 3)。

ジュリア殺害の日、ヘレンへの殺害未遂の日、どちらの場面でもロイロットが犯人として示される。複数の殺害方法の選択肢から、密室であり、外傷や薬物反応がないという事実から、小さい有毒生物が関与していた可能性が残った。また登場人物が、屋敷のなかをくまなく見回るが、当該生物の存在を認識していないことから犯行時や通常時の当該生物の所在は、ロイロットの部屋や金庫内であることも SAT 問題の解として導かれた。

本チャレンジでは、テンソル解析では知識補完を、SAT 問題では整合する状況の探索を試みた。すべての可能な範囲を探索することは非現実的で、人の手で注目する軸や範囲を規定しなければならないという点で、テンソルに 3 軸を与えたことや、SAT 問題でメソッドや限量子を与えたことは、本質的にはフレーム定義とフレーム選定であり、等価であると考えられる。

今回、SAT 問題のアプローチによって説明力を確保し、機械的な推論を行うことを試みた。論理式や事実にも重みをつけて解く Partial Weighted SAT 等に拡張することや、解空間を狭める探索方法を進化計算等で行うことで、論理式(法則)の補完というアルゴリズムの高度化を、SAT 問題の枠組みで行えるようにすることが今後の課題である。我々が今回 SAT 問題解法を利用したことによって、解釈性の高い人工知能・機械学習の基礎技術への探求の一助となることを期待する。

尚、今回の解析には CCG2lambda と AlloyAnalyzer が貢献した。

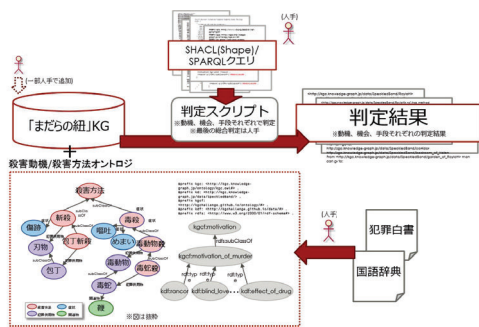


図 4: 全体構成 (上小田中 411)

3.2 富士通研究所 (上小田中 411) の取り組み

3.2.1 全体構成と作成したオントロジー

【全体構成】上小田中 411 チームは動機、機会、方法のそれぞれを持つ人物を推論することで犯人を探し出そうと考えた。

図 4 の様に殺害の動機、方法を示すオントロジーを作成し、提供されたナレッジグラフに追加し、それぞれを判定するスクリプトを適用した。それぞれの判定結果から、最後の総合判定は人手で行った。

【作成したオントロジー】犯人を推理するためには、証拠を積み重ね、犯罪に関連する知識と照らし合わせ、分析を行なう。この今回は a) 動機の分析、b) 機会の分析、c) 手段の分析の 3 つについて知識と推論規則を記述した。犯行の動機については、犯罪白書に犯罪の動機の統計があり、主にそこから列挙した。犯罪の動機の関連情報としてのエージェント間関係は、遺産相続を記述するという観点から The Agent Relationship Ontology に基づいて記述した。手段に関しては、殺人手段を国語辞典の「殺」の後方一致をベースに殺し方を列挙した。またそれぞれの殺し方に対して、国語辞典の語義をベースに手段、供用物などの属性ごとに分離して、詳細を記述した。

3.2.2 動機、機会、方法を持つ人の判定

【動機】殺人の動機を持ってもおかしくない状況かつ対象が明確であれば殺害の動機を持つとし、殺人動機が生じる状況から、該当する人物を推論する規則を作成した。処理系の実行により、1) ロイロットは金欲のためにジュリアとヘレンを 2) 村人は自己防衛のためにロイロットを 3) ヘレンは自己防衛のためにロイロットを殺す可能性がある。という結果を得た。

【機会】事件当夜、ジュリアの部屋に移動可能な場所にいた人物は機会があるとし、事件当時の各々の居場所を推論した。時間のつながりを推論する部分と、空間のつながりを推論する部分に分けた。

事件発生時の登場人物の場所は、事件発生時と同じ時刻を持つシーンの情報を取得し、事件発生以降であることが推論されるシーンを除いた。この推論により事件が起こった時、ジュリア、ヘレン、ロイロットはそれぞれの寝室にいて、ロマは庭にいたと推論された。

次に接続されている空間の繋がりを列挙し、人が通れない繋がりを記述し、差し引いて残ったものを人が移動できる繋がりと推論した。その結果、それぞれの場所からジュリアの寝室に人が通っていくことはできなかつたと推論された。

【方法】事件当夜の被害者や現場の様子から殺害方法を絞り込む部分と、絞り込まれた殺害方法を実行するために必要な条件を満たす人物を導出する部分を記述した。殺害方法を絞り込む部分により、殺害方法は毒殺であり、理由となる症状は「めまい」「青白い」「傷跡が無い」と推論された。人物を導出する部分から、ロイロットに「毒蛇殺」と「毒動物殺」が実行可能であると推論された。その両方とも理由は、鞭がロイロットの部

屋にあったことであった。

3.2.3 総合判定と今後の課題

【総合判定】前節の判定をまとめると、総合的には、本事件はロイロットが金欲のために、毒蛇を用いて、ジュリアを殺したと推論した。

【今後の課題】今回犯人探しに用いた方法は、殺人事件一般に適用可能であると考えている。ただし、オントロジーは、すべての殺害動機や殺害方法について記述したわけではないので、これを充実させることは今後の課題となる。また今回は総合判定を人手で行ったがこれを機械的に判定することも今後の課題である。

3.3 富士通研究所 (FLL-ML) の取り組み

FLL-ML チームは、検索や推論を利用した手法との比較や併用可能性の検討を目的として、機械学習を利用した犯人の予測とその理由の説明に取り組んだ。

犯人の予測のために、「まだらのひも」以外のシャーロックホームズシリーズの短編に出現する登場人物について、対応する単語ベクトルを特徴として“犯人”、“被害者”、“その他”の分類モデルを学習し、「まだらのひも」の登場人物に適用した。モデルの学習には、「まだらのひも」のナレッジグラフに加えて、青空文庫に収録されているホームズの 22 の短編から生成した 14,619 の「単語区切り文」、「登場人物の一覧」および「犯人・被害者の一覧」の 3 種類のリソースを利用した。

また、犯人と判断した理由の説明として、犯人と意味的に近い文の中から、動機・手段を示す文を提示した。動機・手段文の分類には、「まだらのひも」以外の短編に“動機”、“手段”、“その他”のタグ付けを 2,930 文行い、そのデータから学習したモデルを用いた。その際、文の特徴として、文中に出現する単語と対応する単語ベクトルの平均 (文ベクトル) を利用した。犯人と文の意味的な近さは、犯人と対応する単語ベクトルと文ベクトル間の距離とした。動機・手段文出力の際には、犯人に近い文から動機・手段スコアの上位の文をそれぞれ 30 件ずつ提示し、目視によって動機・手段を示す文を判断し、ラベリングを行った。

予測の結果、犯人スコアが最も高い登場人物はロイロットとなり、次点はヘレンであった。ロイロットを犯人とした場合、被害者はジュリアと予測された。犯行の動機は、癪癪あるいは金銭問題であると推定され、犯行手段は特定できなかった。表 1 は動機を示す文として抽出されたものの一例である。

表 1: 抽出された動機文の例

動機	根拠文
癪癪	あるとき、家内で窃盗事件が立て続けに起こりまして、かんしゃくを起こし、現地で雇った執事を殴り殺してしまったそうです
金銭	ただ数エーカーの土地と、築二〇〇年にもなる屋敷だけは残りましたが、それも抵当に入っている始末でございます

犯人の予測結果は、「まだらのひも」の一般的な解釈と相違なかった。犯行動機については、癪癪と金銭問題が抽出されているが、ロイロットの犯行は計画的なものであるため、癪癪は誤りであった。しかしながら、今回の推論チャレンジで、動機として癪癪を考慮したチームは他になく、知識として網羅することが難しい事柄を、事例を元に粗く捉えることができる点は、学習を用いる利点と言える。一方で、本手法のような単純な文の提示では、「まだらのひも」の犯行手段のような複雑な手続きは説明できないため、手続きを説明するためのデータの

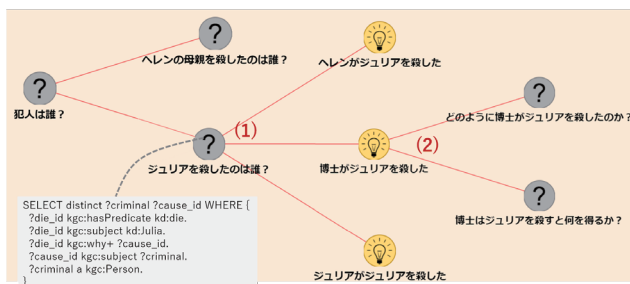


図 5: 誰が犯人か議論するための IBIS 構造

構築および予測結果との対応付けについて検討していく必要があるだろう。

3.4 名古屋工業大学の取り組み

複数のエージェントに誰が犯人か議論させ、構造化された議論のログを使って説明を出力するというアイデアを紹介する。多くの推理小説では探偵役だけでなく、誤った推理を披露する助手役（「サイドキック」等と呼ばれる）を登場させ、複数の仮説について議論する場面が入る。このような場面により、読者が事件の背景や推理の過程を解釈しやすくなる効果があると考えられる。このことから、複数の仮説について自動的に議論させることで、解釈可能性の高い推論過程が実現できると考えた。具体的には、以下のように複数の議論エージェントが意見を戦わせ、IBIS (Issue-based Information System) という構造を用いて「まだらのひも」の真相を推理するというアイデアである。

- 提供されているナレッジグラフに現れる登場人物 x ごとに「 x が犯人」という仮説を立て (図 5(1)), その仮説の説明を試みる議論エージェントを割り当てる。
- ファシリテータエージェントが各議論エージェントに仮説の詳細を質問し (図 5(2)), 各議論エージェントはその質問の答えとなる仮説を生成する。
- 生成された仮説を表す RDF トリプルは、各議論エージェントが保持するナレッジグラフを補完すべく追加される。
- ファシリテータエージェントは各議論エージェントが生成した説明の尤もらしさや整合性を評価し、最終的に評価の高かった説明が出力される。

このような着想に至った背景に、自律的ファシリテータエージェントの研究 [池田 18] がある。IBIS (Issue-based Information System) と呼ばれる議論構造で Web 上の議論を理解し、ファシリテートする研究である。この研究の新しい方向性として、議論参加者にも自律的エージェントが混ざり、人間の参加者が出した意見を補完するような仮説を生成・提示できれば、議論支援に繋がる可能性がある。その題材として、ナレッジグラフ推論チャレンジに挑戦した。なお、本アイデアは未実装であり、仮説をどのようなアルゴリズムによって生成するのは今後の課題である。

4. 評価方法

1 回目となる本チャレンジでは、基本情報の共有の後、専門家評価、一般評価の組み合わせで評価を実施した。但し、評価方法については今後も改善を続けていく予定である。今回、審査員は SWO 研究会の幹事を中心に 7 名に委託した。

専門家による評価では、以下の項目について 1 週間以上の時間を掛けて審査員に評価していただいた。いずれも 5 段階評価 (1-5) である。

- 技術性 (Significance)
- 汎用性 (Applicability)
- 発展性 (Expansibility)
- 知識・データ構築の工夫
- 知識・データ活用の工夫
- 実現可能性 (アイデアのみ)
- 論理的説明性
- 努力性
- 総合的なコメント (自由記述)

一般審査では時間的な制約から説明の心理的な側面 (納得性) に絞って評価を実施した。参加者 45 名が以下の 2 点についてアンケートに答える形で審査を行った。いずれも 5 段階評価 (1-5, 但し, 0.5 刻み) とした。

- 総合評価
- 説明性

一般審査の結果については、平均値では総合評価と説明性いずれも 1 位が 2 位を上回っているが、中央値では総合評価は上回っているが説明性では同値となり、標準偏差では総合評価と説明性いずれも 1 位のほうが大きいという結果となった。一方、専門家による評価で両者を比較すると、各指標の平均では 1 位が 2 位を上回っているが、説明性の観点では逆転した結果が得られた。

そのため、1 位と 2 位の順位付けは専門家審査員による議論に任せられる形となったが、説明性以外の観点ではいずれも 1 位は 2 位と同値、または上回っていたことから今回の結果を決定した。同時に、本チャレンジの一番の課題である説明込みで推論・推定技術の評価する点については、今後の課題となった。尚、その他に各専門家のコメント等を踏まえてベストリソース賞とアイデア賞を授与した。

5. まとめと今後の課題

本論では、説明性、解釈性を有する AI 技術の深耕を目的に開催した第 1 回ナレッジグラフ推論チャレンジ 2018 について報告した。2019 年度も推理小説を対象に、タスクの難易度や傾向を変えながら計 5 編の短編小説からなるナレッジグラフを構築し、2 回目のチャレンジを開催する。更に、推理小説以外にも現実の社会問題や研究データも対象としていきたい。具体的には、社会問題解決に関するベストプラクティス集などを検討している。培われた推論・推定技術を実際の研究課題・社会問題に関するナレッジグラフ等に適用し、説明性を有する AI 技術の社会実装に繋げたい。更に、国際イベント化し、日本初の学術的な問題設定として普及させていくことも検討している。

謝辞

本チャレンジ開催にあたりまして、これまでに開催されたワークショップ、勉強会、発表会、報告会等、各イベントへの参加者の皆様に感謝申し上げます。

参考文献

- [川村 18] 川村 隆浩, 江上 周作, 長野 伸一, 大向 一輝, 森田 武史, 山本 泰智, 古崎 晃司: 第 1 回ナレッジグラフ推論チャレンジ 2018, 第 32 回人工知能学会全国大会論文集, 1F1-01 (2018)
- [池田 18] 池田 雄斗, 白松 俊, 伊藤 孝行, 神谷 晃, 内藤 勝太, 芳野 魁, 鈴木 祥太: Web 議論の自動ファシリテーションのための事前知識を用いた質問生成手法, 人工知能学会 第 83 回言語・音声理解と対話処理研究会, SIG-SLUD-B801-05 (2018)