



8/27(火) Nagatacho GRiD

人工知能学会SWO研究会ワークショップ
「ナレッジグラフ推論チャレンジ2019技術勉強会」

第2回ナレッジグラフ推論チャレンジ2019
応募に向けて

➡ 第1部:ナレッジグラフに関する基礎知識

10:30 ナレッジグラフ入門

- ・知識グラフ, Linked Data, RDF
- ・SPARQLクエリによる検索演習【ハンズオン】

(12:00-13:00 昼食休憩)

➡ 第2部:第2回ナレッジグラフ推論チャレンジ2019応募に向けて

13:00 第2回ナレッジグラフ推論チャレンジ2019

- ・開催概要
- ・ナレッジグラフのスキーマ・構築方針の説明
- ・第2回の対象小説の紹介

14:45 第1回ナレッジグラフ推論チャレンジ2018の振り返り

- ・応募作品の概要
- ・受賞作品の紹介

16:15 ツール部門(第2回から新設)の紹介

- ・想定しているツール例

8/27(火) Nagatacho GRiD

人工知能学会SWO研究会ワークショップ

「ナレッジグラフ推論チャレンジ2019技術勉強会」



KGC: Knowledge Graph Completion Challenge

ナレッジグラフ推論チャレンジ

第2回ナレッジグラフ推論チャレンジ2019 の紹介

◆ シャーロック・ホームズのような

“推理”（推論）ができるAIシステムの開発
を目指した技術コンテスト

➡ チャレンジのねらい

➤ **説明可能性（解釈可能性）**を有するAI技術に関する最新技術の促進・共有と、その分析・評価、体系化を行う。

➡ チャレンジタスク

➤ **推理小説のナレッジグラフ**（ホームズの短編小説）を対象に、ホームズと同じ結論に辿り着き、その理由を説明する。

第2回ナレッジグラフ推論チャレンジ2019

- ➡ 2019年6月4日 応募開始
 - 人工知能学会全国大会(JSAI2019)@新潟にて、第1回開催報告 & 第2回応募開始の宣言
 - 対象とするナレッジグラフについては、調整中のものを仮公開。
- ➡ 2019年8月末 改良版のナレッジグラフを公開
- ➡ 2019年8月27日 技術勉強会
 - ナレッジグラフ・LODの基礎技術から、推論チャレンジのナレッジグラフを扱う技術例まで、ハンズオンを交えた解説。
- ➡ 2019年9月13日 応募相談会
- ➡ 2019年10月末 応募締切
- ➡ 2019年11月末-12月 最終選考会 & 授賞式

- ▶ **説明可能性(解釈可能性)**を有するAI技術に関する最新技術の促進・共有と, その分析・評価, 体系化を行う.
- ▶ 特に, 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, 確率的関係性を含む問題を扱うため, **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技術を対象とする.

- ➡ **推理小説のナレッジグラフ**（シャーロックホームズの短編小説）を対象に，ホームズと同じ結論に辿り着き，その理由を説明する.
 - 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
 - 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，など
- ➡ 但し，第4回チャレンジ以降は社会問題解決に関するベストプラクティス集なども予定

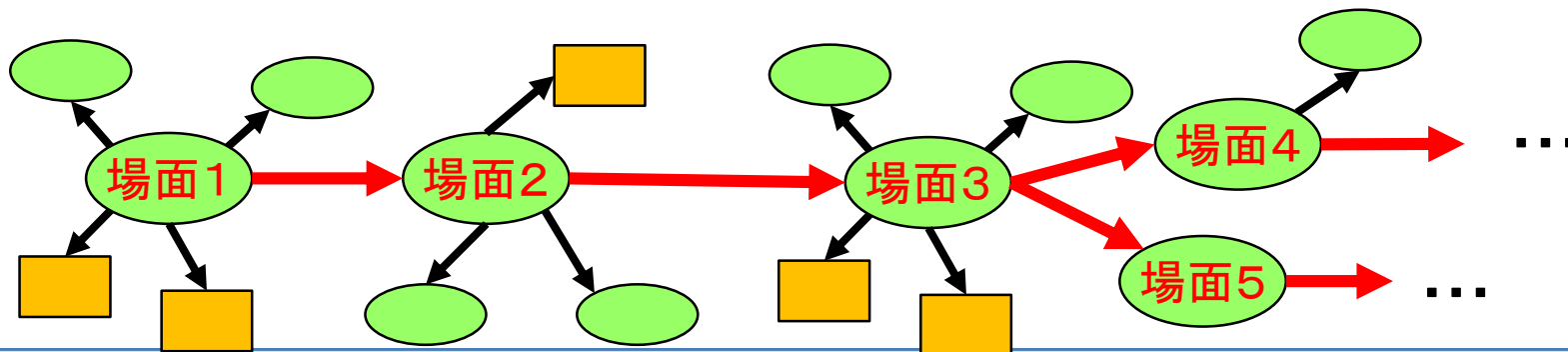


ナレッジグラフの要求仕様

- 犯人を推論(推理)するのに必要な知識を提供する
- 「推理小説」で描かれる様々な状況を, できるだけ統一的な形式で計算機処理(検索・推論・etc.)可能にする

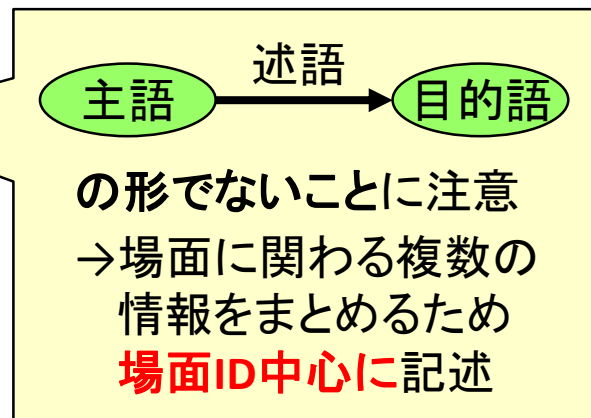
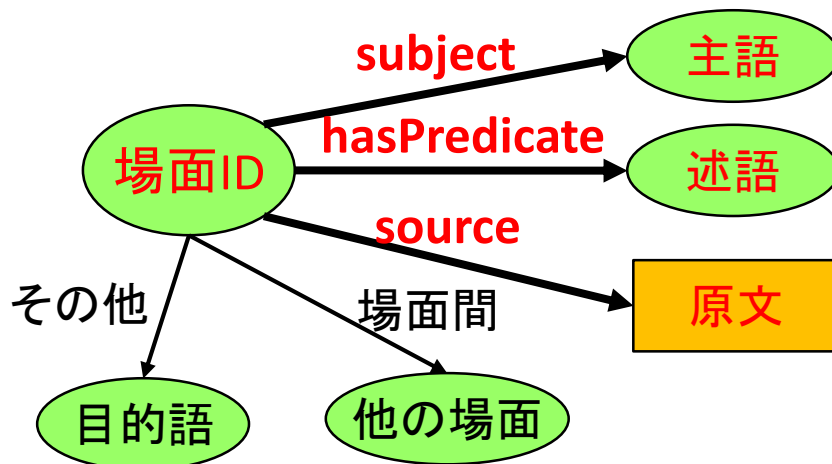
ナレッジグラフ化の基本方針

- 「推理小説」の内容を, 最小単位の「**場面(シーン)**」に分割
→ 場面ごとにID(IRI)を付与
- 「**各場面の記述内容**」および「**場面間の関係**」をグラフ化
→ グラフ化に必要なクラス・プロパティを定義

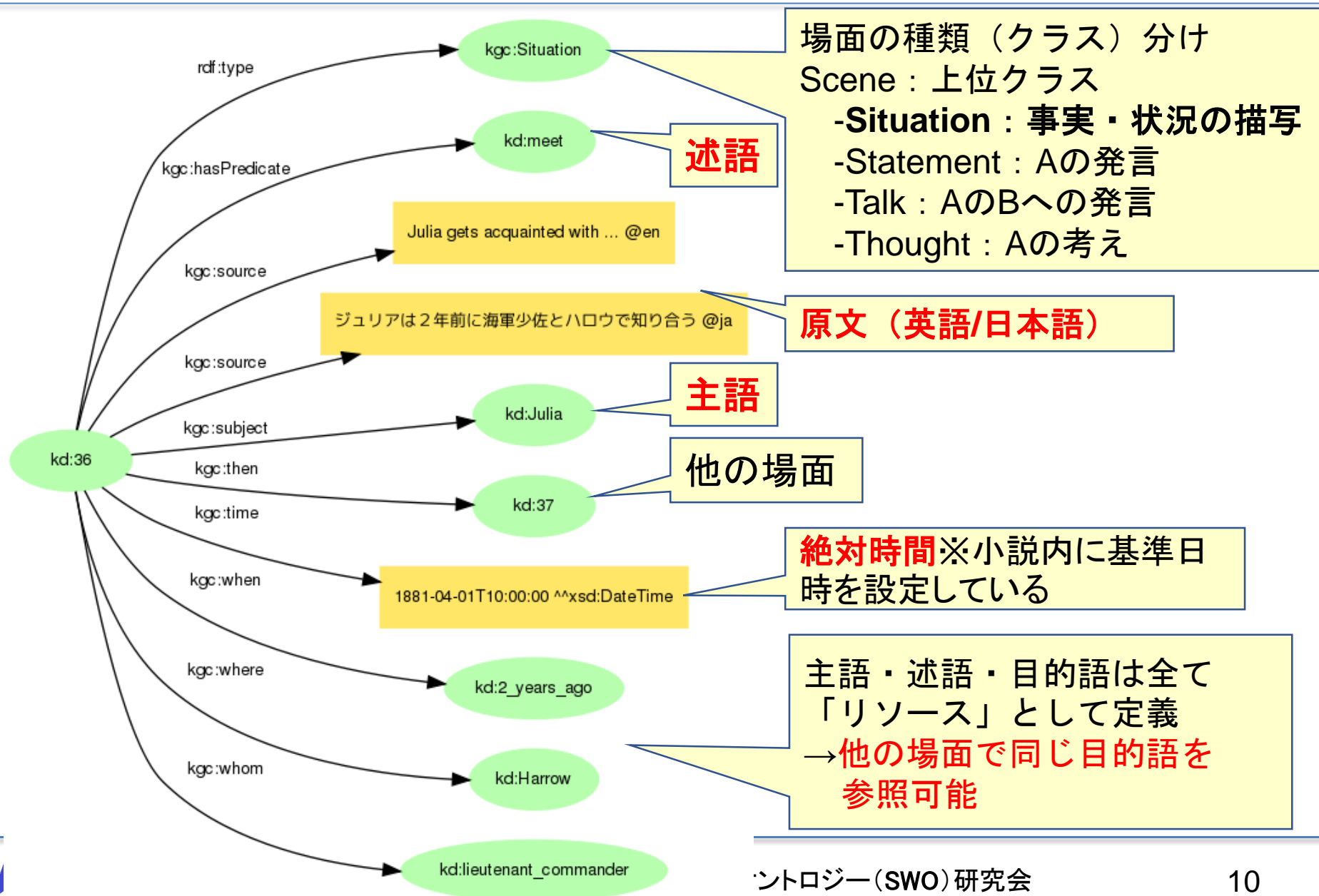


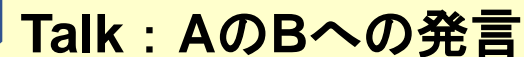
➡ 場面を表現するプロパティ

- **subject**: その場面の記述において主語となる人や物
- **hasPredicate**: その場面の内容を表す述語
- **場面の詳細を表す目的語**: whom(だれに), where(どこで), when(いつ), what(何を), how(どのように), ...etc.
- **場面間の関係**: then, if, because, ...etc.
- **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)



場面(シーン)スキーマ 記述例

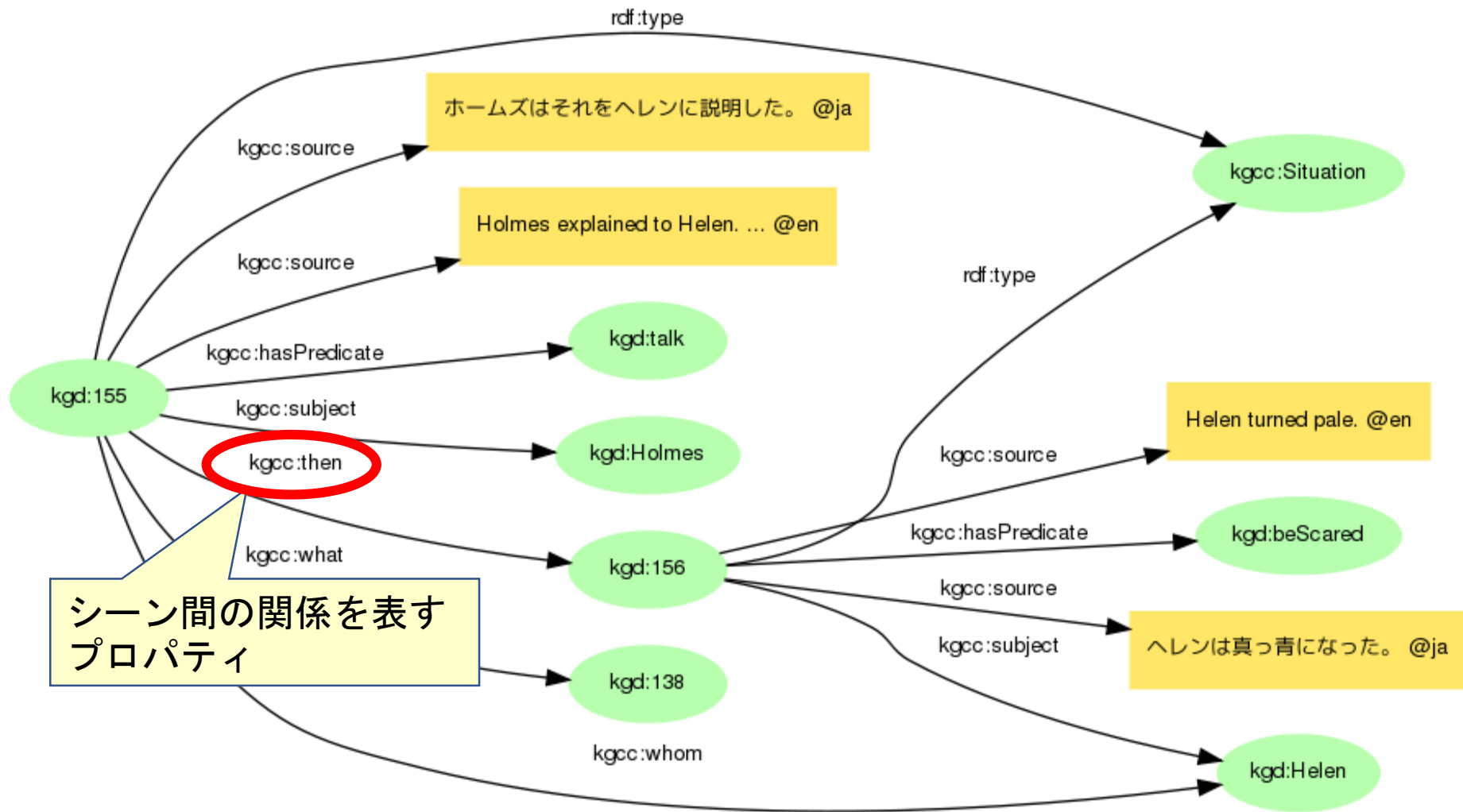


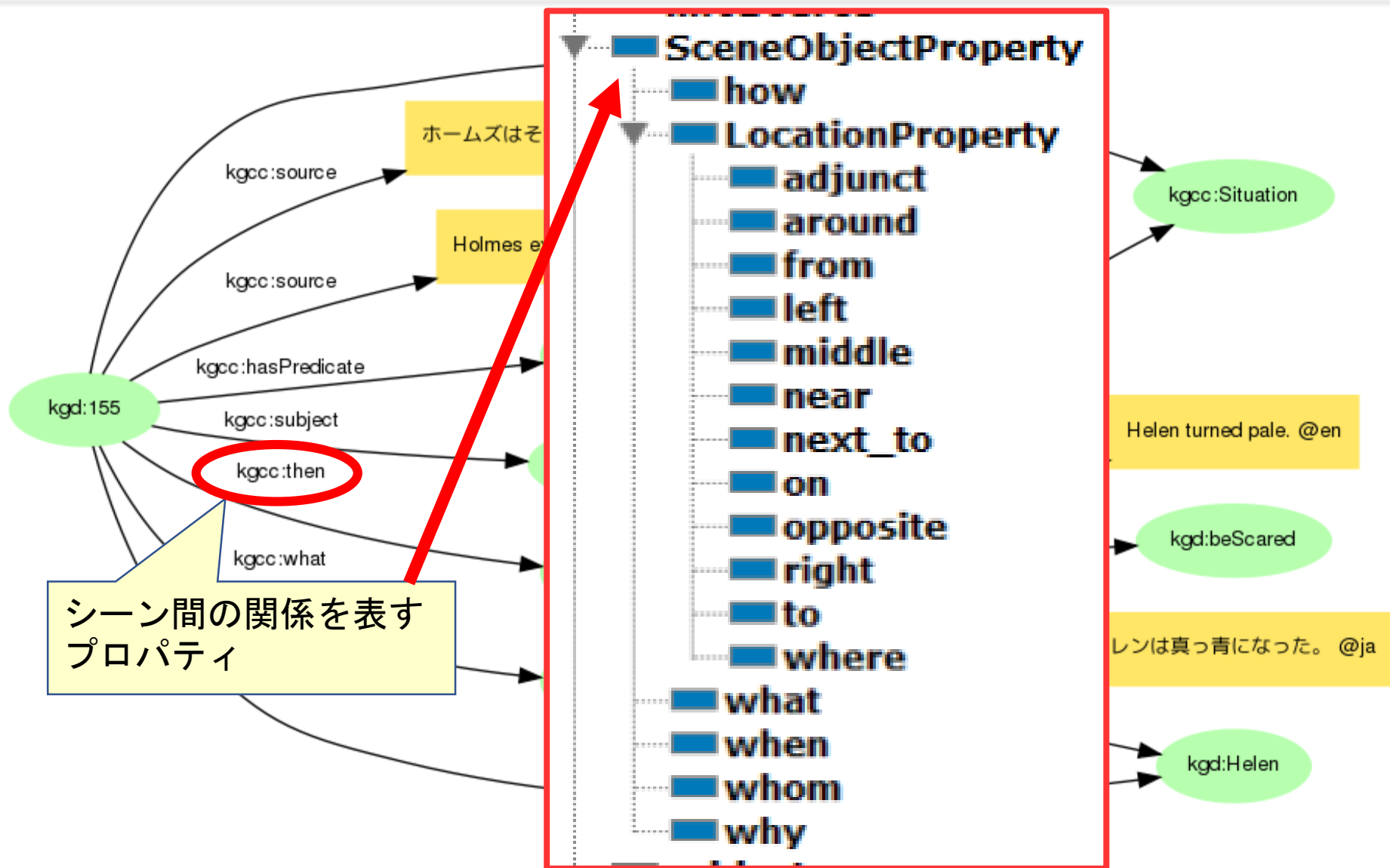


情報源となる人物A
と受け手B

情報源となる人物A

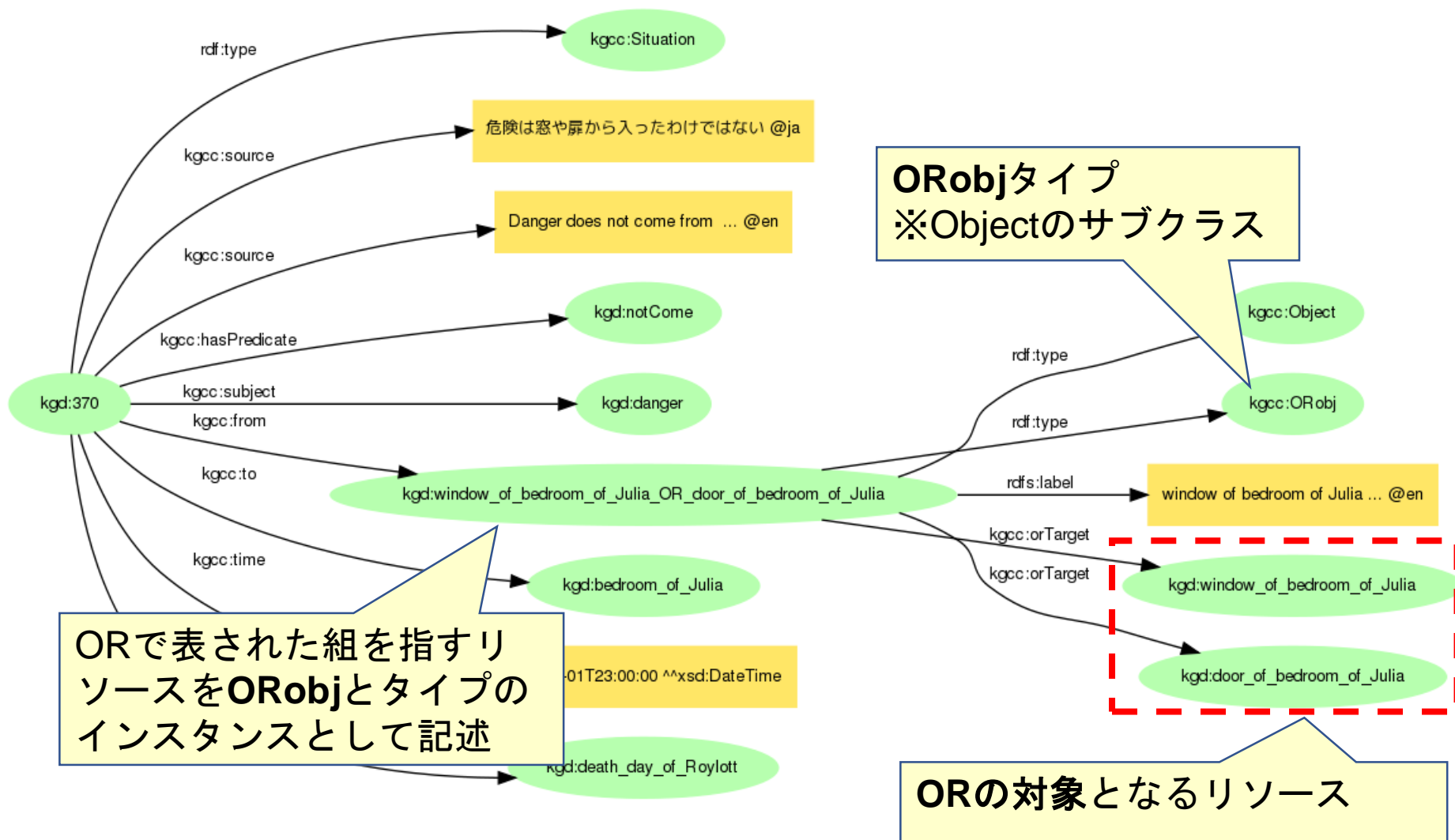
※Talkの場合のみ



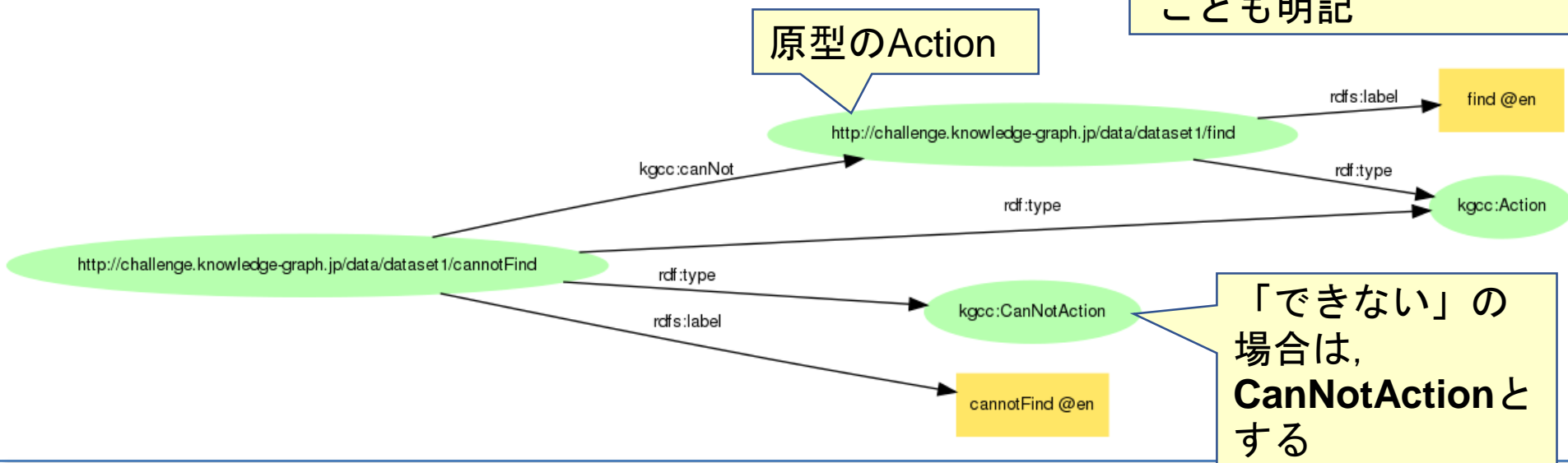
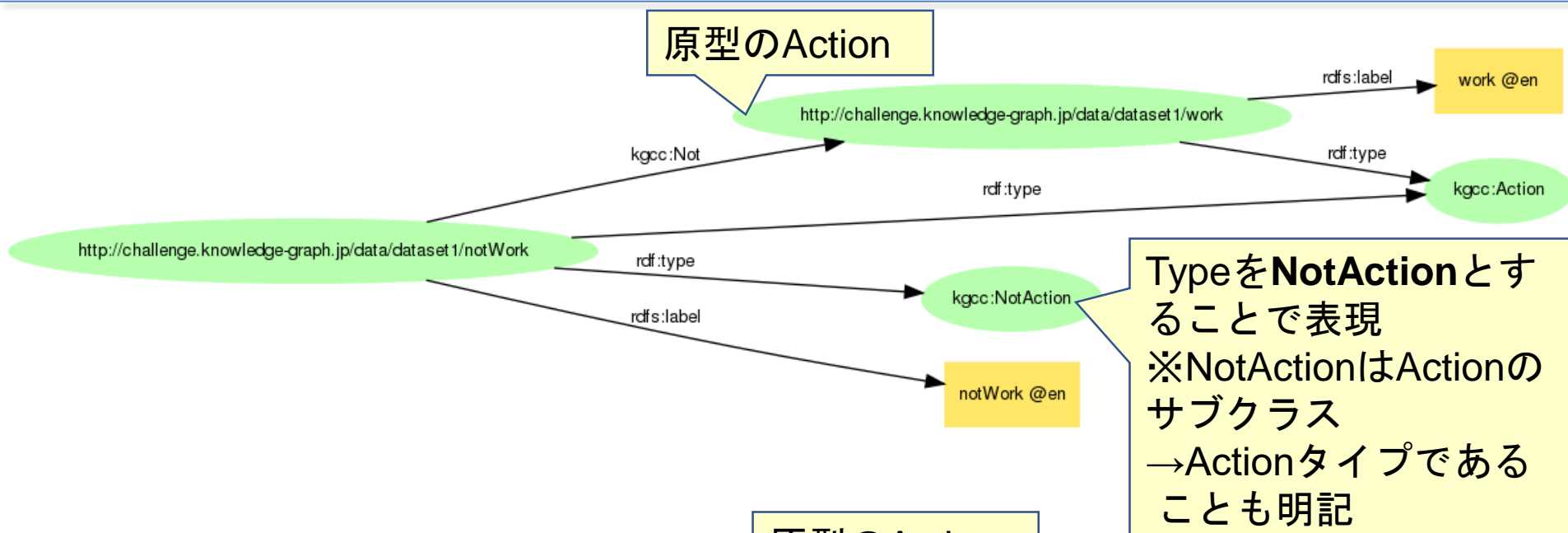


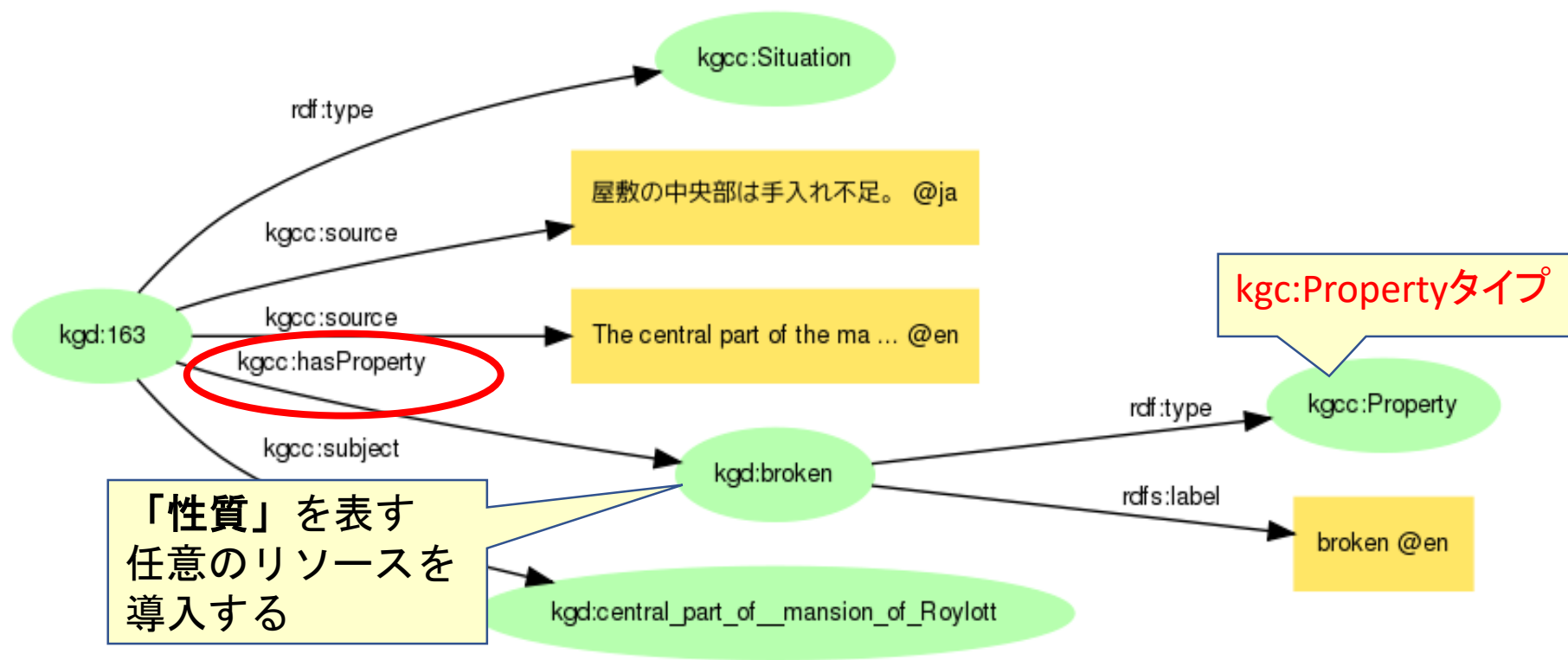
同じプロパティによるトリプルを複数記述することで表現する

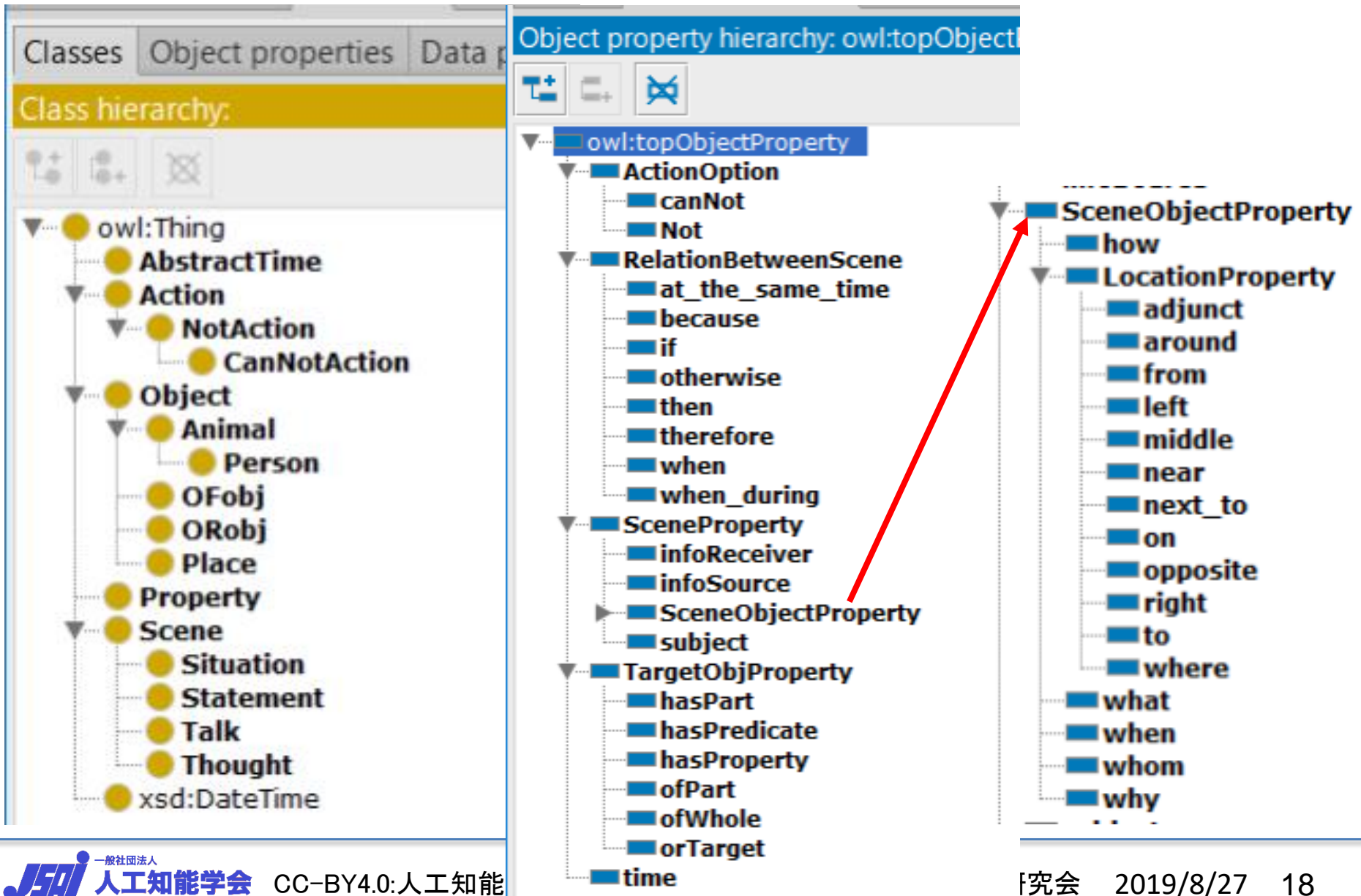




動詞の「Not」の扱い







SWO研究会・勉強会での予備的作業を経て、有志数名でナレッジグラフ化

```

PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX kgc: <http://kgc.knowledge-graph.jp/ontology/kgc.owl#>
PREFIX kd: <http://kgc.knowledge-graph.jp/data/SpeckledBand/>
SELECT ?s ?p ?o
WHERE {
  ?s ?p ?o .
  filter(?p != rdfs:comment)
  filter(?p != kgc:source)
  filter(?o != kgc:Situation)
  filter(?o != kgc:Action)
  filter(?o != kgc:Object)
}
limit 1000
  
```

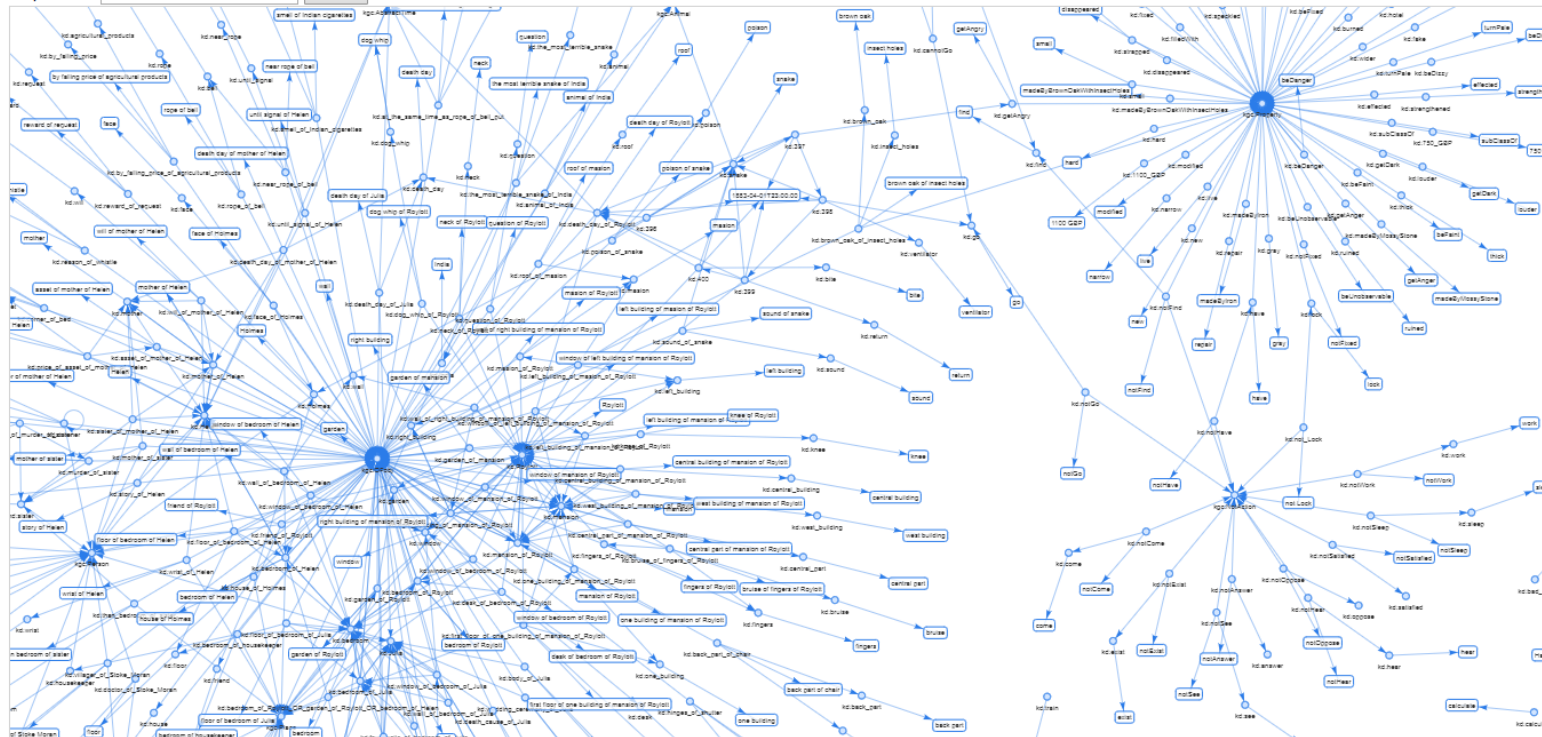
クエリ言語SPARQLによる検索

SPARQL Endpoint: <http://lod.hozo.jp/repositories>

 Keyword:

グラフDB(キーワード検索也可)

ナレッジグラフ(RDF形式)



<http://knowledge-graph.jp/visualization/>

▶ ナレッジグラフの構築方法の改良

- 構築作業の効率化による対象KGの拡大

▶ 対象とする小説の追加

- 1 + 追加4編 = 計5編に

✓ 複数の小説を対象にすることで、より広範囲な課題に取り組めるように！

▶ 「ツール部門」の新設

- タスクを部分的に解く小規模なツールの開発でも応募可能に
- **LODチャレンジとの重複応募もOK！**



➡ ナレッジグラフ(KG)の構築方針

- 基本的に、第1回と同じスキーマを用いて構築する。
- 第1回の開催時に参加から得たフィードバックは適宜反映。
→scene間の関係を表す語彙の検討, など
- 複数KG間での語彙統一など, スキーマの改良(調整中)

➡ 構築作業の効率化の工夫

- 一部の構築作業をアルバイト雇用により実施
→将来的には, クラウドソーシングによる効率化を検討
- 部分的に, (半)自動処理を導入
→将来的には, KG構築タスクのチャレンジ化を検討
- 最終的なKGは, 上記の処理を経たものをベースにして, 運営メンバーで分担して構築

- ➡ 対象とする小説が5つに増えたことで、異なる小説を横断した処理ができるようにスキーマを拡充
- ➡ 各小説ごとに「ベースIRI」および「グラフIRI」を導入
 - <http://kgc.knowledge-graph.jp/data/**小説名**/>
 - 例:「踊る人形」のシーン100は
<http://kgc.knowledge-graph.jp/data/**DancingMen**/100>
- ➡ 小説をまたいで共通化した語彙
 - hasPredicateで参照する「述語」、および、hasPropertyで参照する「性質・状態」
→<http://kgc.knowledge-graph.jp/data/**predicate**/XXX>
 - 固有名詞(例:ホームズ, 地名)
→<http://kgc.knowledge-graph.jp/data/**YYY**>
は, 小説をまたいで, 共通のIRIで定義・参照



▶ 対象とするKG

- 5つの小説のうち、いずれの小説を対象にしてもよい
(どれか1つだけ, 2つだけ…などでもOK)
- できるだけ多くの小説が、同じシステム(仕組み)で解けるとよい
- 各小説で使用するKGの範囲を段階的に変える(昨年同様)
→ 完全(すべてのKG) / 不完全(10%) / 不完全(10%)
- ナレッジグラフの独自拡張も可能(昨年同様)

▶ 対象とするタスク

- ①本部門: 対象小説1つ以上のタスクを解くシステムを開発
- ②ツール部門: いずれかのタスクを部分的に解くツールを開発
例) 容疑者の推定, アリバイ検証, 動機説明, など
★「自然言語文をトリプル化」するKG構築支援ツールの応募も可
- アイデア部門: ①, ②の実現方法のアイデア(実装なしでOK)

➡ 第1部:ナレッジグラフに関する基礎知識

10:30 ナレッジグラフ入門

- ・知識グラフ, Linked Data, RDF
- ・SPARQLクエリによる検索演習【ハンズオン】

(12:00-13:00 昼食休憩)

➡ 第2部:第2回ナレッジグラフ推論チャレンジ2019応募に向けて

13:00 第2回ナレッジグラフ推論チャレンジ2019

- ・開催概要
- ・ナレッジグラフのスキーマ・構築方針の説明
- ・第2回の対象小説の紹介

14:45 第1回ナレッジグラフ推論チャレンジ2018の振り返り

- ・応募作品の概要
- ・受賞作品の紹介

16:15 ツール部門(第2回から新設)の紹介

- ・想定しているツール例



8/27(火) Nagatacho GRiD

人工知能学会SWO研究会ワークショップ
「ナレッジグラフ推論チャレンジ2019技術勉強会」

第2回ナレッジグラフ推論チャレンジ2019
の対象小説の紹介



対象とする推理小説

➡ まだらのひも(第1回のKGの不具合を修正して利用)

- タスク: ヘレンを殺したのは誰か? (犯人+説明)

に加え, 新たに, 以下の4編をKG化

➡ 踊る人形 [[Wikipedia](#)] [[青空文庫](#)]

- タスク: 暗号を解け (暗号の解読)

➡ 背中の曲がった男(曲がれる者) [[Wikipedia](#)] [[青空文庫](#)]

- タスク: バークリはなぜ死んだのか? (説明)

➡ 悪魔の足 [[Wikipedia](#)] [[青空文庫](#)]

- タスク: 各人物を殺したのは誰か? (犯人+説明)

➡ 花婿失踪事件(同一事件) [[Wikipedia](#)] [[青空文庫](#)]

- タスク: 花婿はなぜ消えたか? (説明)

➡ ナレッジグラフの公開場所

- 「推論チャレンジ」のサイト
→ ナレッジグラフの公開

<https://github.com/KnowledgeGraphJapan/Challenge/tree/master/rdf/2019>

➡ 公開形式

- 小説ごとのRDFファイル(Turtle形式)のダウングレード
※他の形式は準備中....

- SPARQLエンドポイント

<http://lod.hozo.jp/repositories/kgc2019>

- 可視化ツール

<http://knowledge-graph.jp/visualization/>

- ➡ 第2回ナレッジグラフ推論チャレンジが対象とする小説
(**新規追加の4篇**)について、応募の参考情報として、

- 小説のあらすじ

- 必要と思われる推論／検索方針の一例

を紹介する。

- ➡ あくまでも、考えるきっかけとするための「一例」なので、
ここで示す例が、「正解例」や「最善の方法」とは限らないことに注意。

- ➡ 示した例を、参考にした応募もOK！



KGC: Knowledge Graph Completion Challenge

ナレッジグラフ推論チャレンジ

第2回ナレッジグラフ推論チャレンジ2019勉強会 ～ 踊る人形 *THE ADVENTURE OF THE DANCING MEN* ～

超簡単なあらすじ紹介（1／2）

➡ 依頼人がホームズに「踊る人形」 を見せる

- 依頼人キュービットは、最近結婚した旧家の旦那である
- キュービット家の近くに何度か落書きされていた
- 奥さんエルシイは何か知ってそうだが過去を語らない
- ホームズ(ウキウキで)調べ始める

➡ 依頼人がホームズに新しく見つけた「踊る人形」を送る

- それを読んだホームズは急いでキュービット家に向かう
- キュービット家では、キュービットとエルシイが撃たれていた
- 警部登場
- ホームズが現場検証と聞き取りを行う

超簡単なあらすじ紹介（2／2）

- ➡ ホームズが新しい「踊る人形」を書いて近所エルリッジに送る
 - ホームズが「踊る人形」の読み方をワトソン、警部に説明する
 - ホームズが自分で書いた「踊る人形」の意味を説明する
- ➡ 訪ねてきた男をホームズと警部が捕まえる
 - 男はアメリカ人のギャングであるスレイニであった
 - スレイニ はペラペラと経緯を自供する
 - エルシイと婚約していたが、逃げられたので追ってきた
 - キュービットに撃たれたので、スレイニがキュービットを撃ち殺した
 - 奥さんは余生を慈善事業に費やした

犯人の推理例（1／2）

①「踊る人形」を解読する

「踊る人形」解読の手がかり

- アルファベットでは E がもっともよく使われる
- 次に使われるのは T, A, O, I である
- 単語の区切りが何かあるはず
- 頻出するパターンを特定し、単語を当てはめる
- 人名の前は呼びかけ(来い, など)である可能性がある
- あとは当てはめて意味の通るものを探す
- ...



人形1つ1つを画像認識して、同じ人形を同じ文字としてシンボル化し、
上記ルールを適用しながら、アルファベットを当てはめて文を解読

解読された「踊る人形」の一覧

		
AM HERE ABE SLANEY	AT ELRIGES	COME ELSIE
		
NEVER	ELSIE PREPARE TO MEET THY GOD	
		
COME HERE AT ONCE		

世界の文字：踊る人形文字

http://www.chikyukotobamura.org/muse/wr_fiction_19.html

犯人の推理方法（2／2）

2. 犯人を推論する

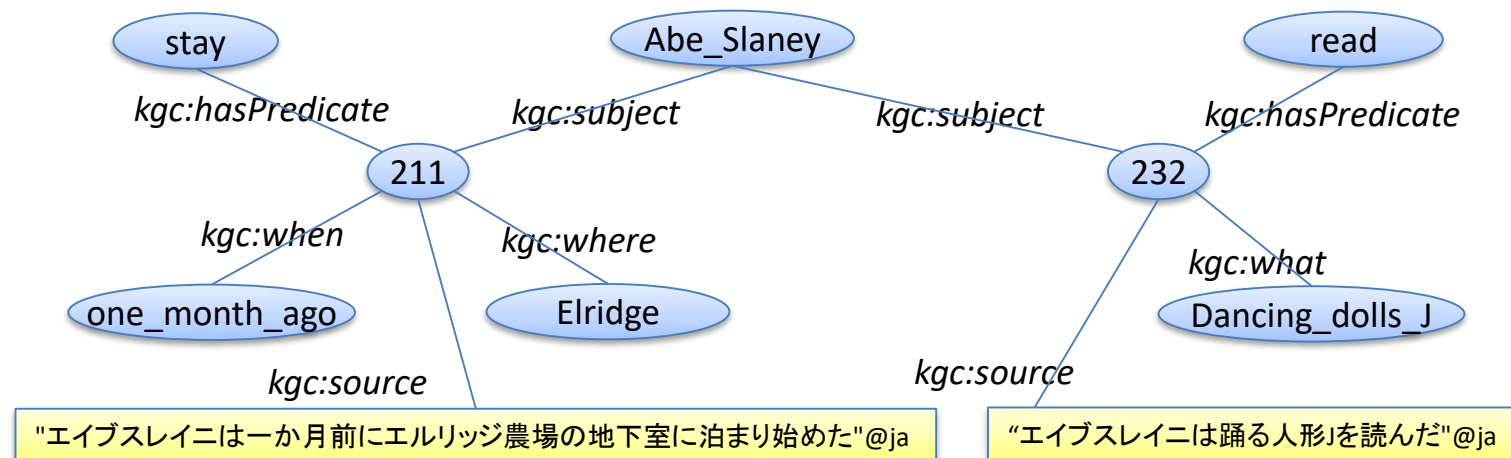
- 「踊る人形」が解読できても犯人の名前は出てこない...
- そこで、1ステップだけのルールを設ける

エルリッジに居る人 $\langle x, \text{stay}, \text{Elridge} \rangle \rightarrow$ 犯人 $\langle x, \text{equalTo}, \text{criminal} \rangle$

あるいは、

踊る人形を読める人 $\langle x, \text{read}, \text{Dancing doll J} \rangle \rightarrow$ 犯人 $\langle x, \text{equalTo}, \text{criminal} \rangle$

- これらを下のナレッジグラフに適用することによって、ホームズの書いた踊る人形を読んで、エルリッジからやってきた人 = スレイニが犯人と推論できる





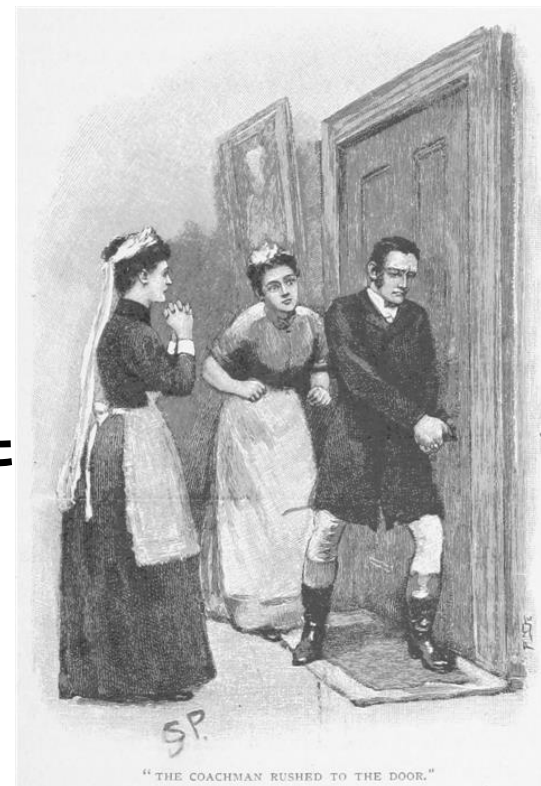
KGC: Knowledge Graph Completion Challenge

ナレッジグラフ推論チャレンジ

第2回ナレッジグラフ推論チャレンジ2019勉強会 ～ 背中曲がった男 THE ADVENTURE OF CROOKED MAN ～

超簡単なあらすじ紹介（1／3）

- ➡ バークリ大佐が殺された件についてホームズが調べているところから始まる
- ➡ 事件はバークリ夫妻と女中が住む邸宅で起きた
 - 夫人が21:15過ぎに聖ジョージ組合の会合から帰宅
 - 夫人は女中に紅茶を一杯頼んだ
 - バークリは食堂にいたが、妻に会いに居間へ向かった
 - 10分後に女中が紅茶を運びに行くと、主人夫婦が言い合うのが聞こえた
 - ✓ 夫人は何度も「卑怯者」「デイヴィド」と叫んだ
 - 夫人の悲鳴が聞こえ、御者が部屋に入る
 - 夫人は気を失っており、バークリは血溜まりの中で死亡していた



曲れる者,アーサー・コナン・ドイル,
三上於菟吉訳, 大久保ゆう改訳
https://www.aozora.gr.jp/cards/000009/files/54911_47220.html

超簡単なあらすじ紹介（2／3）

➡ 現場の状況

- バークリの後頭部には鈍器による裂傷
- バークリの顔はすさまじい恐怖の形相（極度に歪んだ顔）
- 死体のそばに骨の柄がついた木彫りの棍棒があった
- 小動物の足跡（足の肉球が五つ、爪は長い）
- 窓の鍵がなくなっている

➡ 夫人と会合に同行していたモリソン嬢が語る

- 夫人は会合の帰りに猫背の男に声をかけられ驚く
- 男は30年前に死んだと思っていた、ヘンリだった
- 夫人とヘンリが話し終わると、二人はとても怒っていた
- 夫人はこのことを話さないようにモリソンに口止めした
- 夫人が犯人と疑われているので、モリソンは正直にこのことを語った



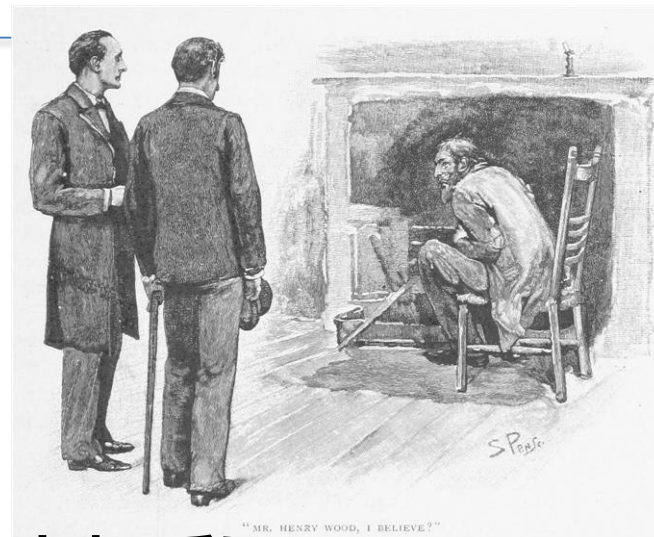
超簡単なあらすじ紹介（3／3）

➡ ホームズとワトソンがヘンリを訪ねる

- 夫人が犯人だと疑われているとヘンリに告げる
- ヘンリは、夫人と自分は無実であると言う

➡ 事件の真相

- ヘンリは30年前バークリと同じ部隊にいた
- ヘンリと夫人は当時愛し合っていたが、バークリも夫人を愛していた
- バークリはヘンリを敵に売ってヘンリは奴隷となり、バークリは夫人と結婚
- この事実を夫人が知り、バークリを「卑怯者！」と罵った
- バークリは言い争いに割って入ったヘンリを見るなり倒れ、頭を打った
- ヘンリは助けを呼ぼうと夫人から戸の鍵を取って中に入ろうとするも、やはり何もせず逃げたほうが得と思って逃げた
- その時、杖を落とし、飼っていたマングースもカーテンを駆け上がってしまう
- 検死の結果、バークリは頭を打ち付ける前に、すでに卒中で亡くなっていた

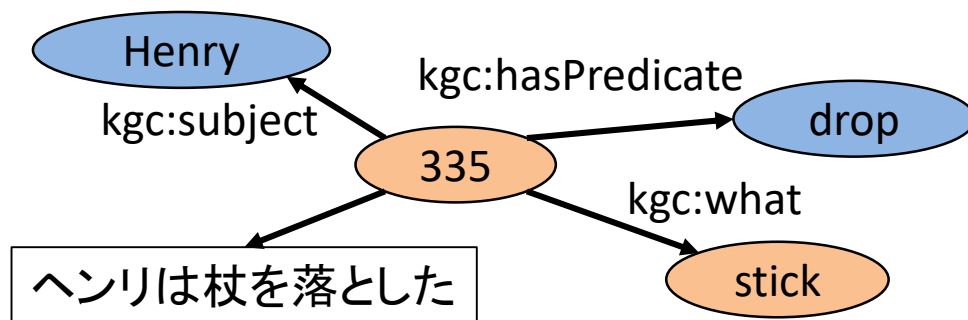
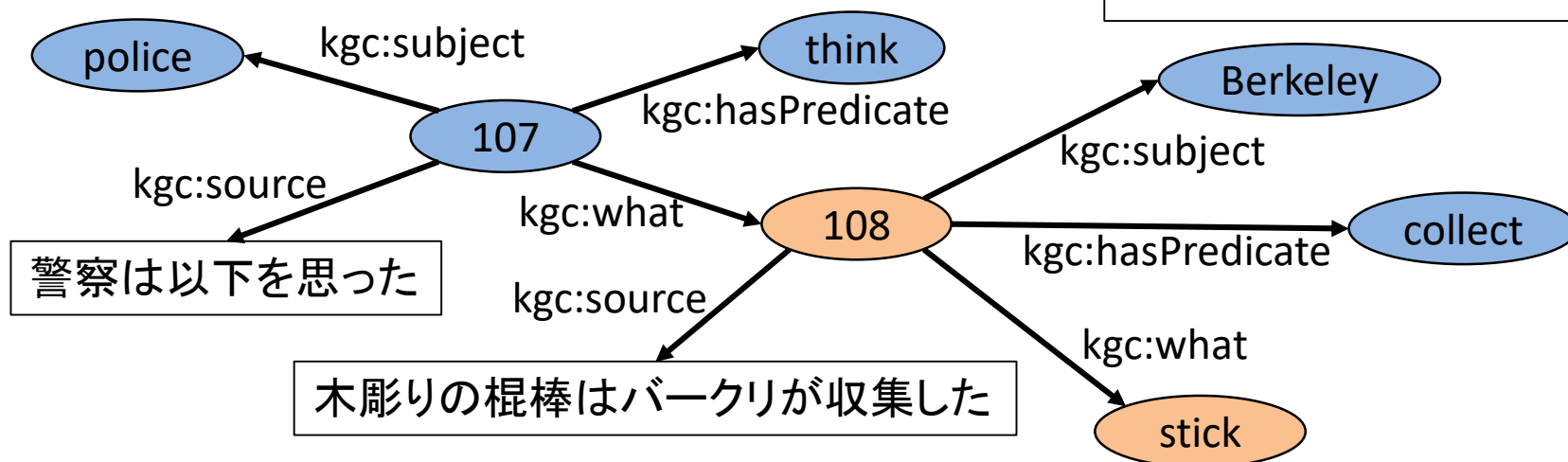


真相の推理例（1／2）

夫人の容疑を晴らす

- 落ちていた棍棒に関するトリプルをSPARQLで取得し、得られたトリプルを辿ってみる

```
SELECT DISTINCT * WHERE {  
  ?s kgc:what ?o .  
  ?o rdfs:label ?str .  
  FILTER REGEX(?str, "stick", "i")  
}
```



持ち主はバークリ
またはヘンリ。
夫人と棍棒に関連
は見当たらない

真相の推理例（2／2）

▶ 夫人の容疑を晴らす(2)

➤ ホームズの仮説

「バークリは男に恐怖して卒倒した」

「バークリは頭を炉格子に打ったかもしれない」
で考える

➤ 考えられるアプローチ

✓ 男(ヘンリ)に恐怖した理由の説明

- 夫人が気絶間際に叫んだ単語からの推測(?)
- 男の容姿から推測(?)

✓ 極度に顔が歪むケースの知識の追加(?)

✓ 自殺, 事故, 病気の考慮



KGC: Knowledge Graph Completion Challenge

ナレッジグラフ推論チャレンジ

第2回ナレッジグラフ推論チャレンジ2019勉強会 ～ 悪魔の足 *The Adventure of the Devil's Foot* ～

超簡単なあらすじ紹介（1／3）

- ▶ ホームズとワトソンが滞在している別荘に、モーティマーが、恐ろしい知らせを持ってくる。
 - 妹のブレンダが死亡し、兄弟のオーウェンとジョージが発狂した。
 - 現場は犠牲者3人の住む家。今朝、死亡・発狂している3人が発見された。
 - モーティマーは、悪魔の仕業ではないかという。
- ▶ 現場で3人の家政婦に事情を聞く。
 - 昨晩は何の異常もなく、今朝3人の姿を居間で発見して自分も恐怖で気絶してしまったという。
 - 現場には侵入の形跡はなく、居間の暖炉には火を入れていたと考えられる。
- ▶ 別荘に戻ると、近所の住人でアフリカ探検家のスターンデイル博士が訪ねてくる。
 - 博士はアフリカに向かう途中に、事件を知らせる電報を受けて戻ってきた。
 - 博士はトリゲニス一家とは親戚で、捜査の状況を聞きたいという。
 - ホームズは博士を尾行する。

超簡単なあらすじ紹介（2／3）

- ➡ 翌朝、モーティマーが自室で前の事件同様に死んでいる。
 - 現場はランプが燻ぶっているせいか、空気が淀み息苦しい。
 - モーティマーは妹のブレンダと同じく、恐怖の表情を浮かべて死んでいた。
 - 現場を調べたホームズは、ランプから灰を採取する。
- ➡ ホームズは、
 - 2件とも閉め切られた部屋で火の存在があり
 - 立ち入った者が気絶したり息苦しくなったりしているから、何かが燃やされ有毒な気体が発生していたと推理する。
- ➡ ホームズは2件目のランプから燃え残りの粉末を採取し、その効果を実際に燃やし確認する。
 - 粉末の燃える臭いを吸い込んだワトソン、恐怖に襲われ苦しむ。
 - ワトソンは、力を取り戻しホームズを抱きかかえて部屋から脱出する。

超簡単なあらすじ紹介（3／3）

➡ 事件の真相

- モーティマーは、博士の家から「悪魔の足の根」というアフリカに伝わる毒草の粉末を盗み出し、博士が出発するのを待って使用した。
- 博士とブrendaは長年の恋人同士であり、この毒が現在の科学では検出できないため、博士は自らの手で復讐すると決める。
- そしてモーティマーのもとを訪れ、同様に「悪魔の足の根」を用いて殺害した。

真相の推理の方針例

- ➡ 2つの事件で，“燃やされた何か？”を推論する必要がある。
 - 事件の場面によって，“灰”，“粉末”，“薬品”など異なる表現がなされているため，それらが同一であることを判定する説明が必要。
 - ✓ 何らかの外部知識を導入？
 - ✓ この小説のナレッジグラフ自体で，そのような知識につながる記述が必要？
- ➡ ナレッジグラフ自体の改変については，引き続き検討予定．提案があれば，ご連絡下さい．