Datasets of Mystery Stories for Knowledge Graph Reasoning Challenge

Kouji Kozaki[1], Shusaku Egami[2], Kyoumoto Matsushita[3], Takanori Ugai[3,2], Takahiro Kawamura[4,2] and Ken Fukuda[2]

[1]Osaka Electro-Communication University, Osaka, Japan
[2]National Institute of Advanced Industrial Science and Technology (AIST), Tokyo, Japan
[3]Fujitsu Limited, Kanagawa, Japan
[4]National Agriculture and Food Research Organization, Ibaraki, Japan

Abstract
With the increasing application of AI systems across various social domains, the explanation for the AI decision is becoming important to guarantee the security and safety of such AI systems. Therefore, the Special Interest Group on Semantic Web and Ontology of JSAI started the Knowledge Graph Reasoning Challenge in 2018. It calls for techniques for reasoning and/or estimating criminals with a reasonable explanation based on knowledge graphs representing well-known stories of Sherlock Holmes. The challenges were held four times in Japan and once as an international event. Through them, 35 works were submitted in total. In organizing the challenge, we have been developing knowledge graphs about eight Holmes mystery stories. They have been extended and improved through the challenges. This paper reports on the knowledge graphs of mystery stories developed and published for the Knowledge Graph Reasoning Challenge.

Keywords
knowledge graph, knowledge modeling for story, dataset, reasoning, explainable AI

1. Introduction

Semantic reasoning using knowledge graphs is expected to increase the explainability of AI systems and contribute to more appropriate use of AI technologies. Therefore, the Special Interest Group on Semantic Web and Ontology of the Japanese Society for Artificial Intelligence (SIG-SWO of JSAI) started the Knowledge Graph Reasoning Challenge in 2018. Then, the challenges have been held annually, five times to date [1, 2, 3, 4]. In the challenge, we constructed and published knowledge graphs based on the Sherlock Holmes mystery stories. Its task is to correctly identify the culprit and causes of incidents and accidents using some AI techniques with the published knowledge graphs.

The reasons for choosing mystery stories as the subject matter include:

SEMMES'23: Semantic Methods for Events and Stories co-located with 20th Extended Semantic Web Conference (ESWC2023), May 29, 2023, Hersonissos, Greece
kozaki@osakac.ac.jp (K. Kozaki); s-egami@aist.go.jp (S. Egami); ugai@fujitsu.com (T. Ugai); ken.fukuda@aist.go.jp (K. Fukuda)
0000-0003-4578-4980 (K. Kozaki); 0000-0002-3821-6507 (S. Egami); 0000-0001-5245-9719 (T. Ugai); 0000-0001-7366-1094 (K. Fukuda)

- They can allow for the design of tasks that are virtually closed (e.g., which have answers and can control the constraints that lead to them) while including complex relationships in the real world.
- Some tasks can be solved without including probabilistic processing or machine learning, such as uncertain information or photographic evidence, or without supplementing common knowledge that is not written explicitly, thus encouraging the fusion of estimation and inference.
- They have an explanatory quality to human beings that the reader must agree with in order for it to work as a novel.
- The stories are widely known to the public and easily attract interest.

A total of 35 entries have been submitted through the five challenges held to date. The techniques used in these works include search and inference based on knowledge processing using knowledge graphs and ontologies, machine learning using knowledge graph embedding, and natural processing techniques, and so on.

In organizing the challenge, we have been developing knowledge graphs eight Holmes mystery stories in total. They have been extended and improved through the challenges. We believe that these knowledge graphs are not only the target data for the challenge, but also a valuable dataset for reasoning and explanation tasks using knowledge graphs. In this paper, we report on the knowledge graphs of mystery stories developed and published for the Knowledge Graph Reasoning Challenge.

In the following sections of this paper, Section 2, we present an overview of the technical challenges we have faced so far. Section 3 describes the schema design and the knowledge graphs published for the of challenge. Section 4 lists the related works on event-centered knowledge graphs, and finally, in Section 5, we summarize the results and discuss future challenges.

2. Knowledge Graph Reasoning Challenge

2.1. Purpose of the Challenge

The Knowledge Graph Reasoning Challenge is a technical contest organized by the Special Interest Group on Semantic Web and Ontology of the Japanese Society for Artificial Intelligence (SIG-SWO of JSAI). The contest was launched in response to the growing interest in AI technologies, particularly deep learning, and the associated emergence of issues related to the explainability of AI systems. It calls for techniques for reasoning and/or estimating criminals with a reasonable explanation based on knowledge graphs representing well-known stories of Sherlock Holmes.

The task for the challenge is to correctly identify the culprit and causes of incidents using inference and estimation techniques. However, since it can be generalized as a kind of knowledge graph completion [1], it can be positioned as a generic problem setting that can be applied

[1]For example, we can regard it is a completion task to find any missing relationship between a particular person and the culprit.

to the construction of various knowledge bases including knowledge graphs, information extraction and relation extraction, knowledge updating and maintenance, and so on. Moreover, in addition to the focus on real social problems and the emphasis on the explainability of the results, there are some unique difficulties, such as described in the following points:

- Real-world problems are all individual cases, and similar scenes do not necessarily appear more than once. Therefore, knowledge or data is not necessarily big data, making learning difficult.
- Rather than explaining single relationships by approximation in vector space, they must be assembled or chained together to derive the goal as a whole.
- The knowledge graph includes false statements spoken by the characters.

2.2. Target Stories and Tasks

In this study, we provided the contents of eight of Sherlock Holmes' s short mystery stories as knowledge graphs in total with tasks which should be solved using them. The target stories that provided the knowledge graphs and their respective tasks are listed below.

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

The contents of these eight mystery stories were converted into knowledge graphs based on events and scenes. Participants in the technical challenge developed AI systems using these data together with their own external knowledge, which was added and created as necessary. The knowledge graph and past proposed techniques are available on the official website[2].

Details of the knowledge graphs and some examples are given in Section 3.

2.3. History

We held the 1st Japan domestic knowledge graph reasoning challenge in 2018. Then, we had five challenges in total every year. Four of them are Japan domestic version, and the other is international challenge.

There were three categories for application to the challenge, as follows

1. Main track: develop a system to solve one or more tasks of the target stories.
2. Tool track: develop tools to solve partially any of the tasks (e.g., suspect estimation, alibi verification, motive explanation, and so forth).
3. Idea track: derive ideas on how to realize any of the above (possibly without system implementation).

[2]https://ikgrc.org/2023/

The total number of proposals was 35 (16 in the main track, 9 in the tools track, and 10 in the ideas track). The following is a summary of the knowledge graph reasoning challenge held so far, along with the number of submitted works:

- 1st Knowledge Graph Reasoning Challenge 2018:
  - Provided KG: The Speckled Band
  - Number of submitted works: 8 (Main Track: 5, Idea Track: 3)
- 2nd Knowledge Graph Reasoning Challenge 2019:
  - Difference from the previous year: Introduction of the Tool Track (solving partial tasks) and addition of four KGs
  - Provided KGs: The Speckled Band, A Case of Identity, The Crooked Man, The Dancing Men, and The Devil's Foot
  - Number of submitted works: 9 (Main Track: 4, Tool Track: 2, Idea Track: 3)
- 3rd Knowledge Graph Reasoning Challenge 2020:
  - Difference from the previous year: Refinement of the existing KGs and addition of three KGs
  - Provided KGs: The Speckled Band, A Case of Identity, The Crooked Man, The Dancing Men, The Devil's Foot, The Abbey Grange, Silver Blaze, and The Resident Patient
  - Number of Submitted Works: 7 (Main Track: 3, Tool Track: 2, Idea Track: 2)
- 1st Knowledge Graph Reasoning Challenge for Students 2021:
  - Difference from the previous year: Applicants were limited to students to foster young talent. In addition, the existing KGs were refined.
  - Provided KGs: same as the previous year
  - Number of Submitted Works: 5 (Main Track: 2, Tool Track: 3)
- 1st International Knowledge Graph Reasoning Challenge 2023 (IKGRC2023):
  - The challenge was internationalized. In addition, the existing KGs were refined.
  - Provided KGs: same as the previous year
  - Number of Submitted Works: 6 (Main Track: 2, Tool Track: 2, Idea Track: 2)

3. Knowledge Graphs of Mystery Stories

3.1. Knowledge graph schema

We decided on a basic policy of describing the people, things, and places involved in each scene, focusing on the scenes depicted in the scenes and the relationships between the scenes. When designing the schema, in addition to expressiveness to represent the subject novels, we also considered the ease of constructing the knowledge graph and the convenience of providing it as data for inference processing, and decided on a schema with mainly 5W1H edges, focusing on scenes. Thus, a mystery story is represented by each scene and the relationships among

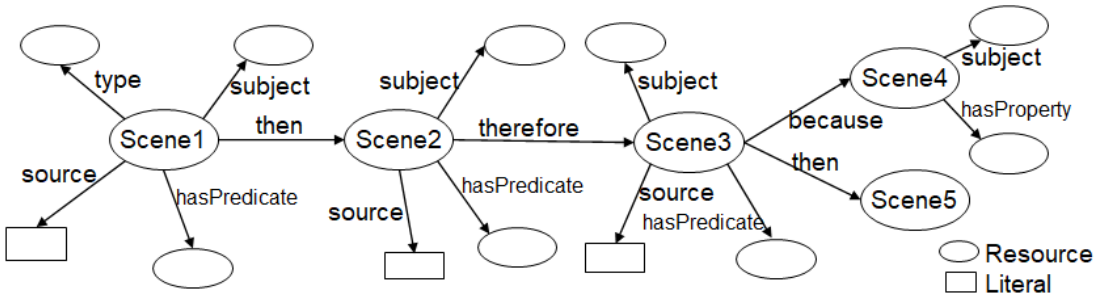

Figure 1: Scene knowledge graph.

scenes. Each scene[3] in a mystery story is assigned a unique internationalized resource identifier (IRI), which is used as the subject to describe a scene in the story by adding information about people, organizations, and places as objects. The relationships between scenes explain the causal relationships of chronological actions and events by referring to the IRIs. This is how a series of storylines is expressed. In addition, rules and table data can be linked to describe common sense data such as axioms and to represent information such as timetables. The content of the story is stored as literal values for natural language processing. Figure 1 shows an example knowledge graph.

The following basic properties are provided for describing each scene. In order to summarize the information associated with a scene, these properties take the scene as their subject. Note that it is not in the general <subject, predicate, object> format. Figure 2 shows an example scene description.

- subject: a person or thing that is the subject in the description of the scene.
- hasPredicate: a predicate that describes the content of the scene.
- hasProperty: a property of the person or thing that is the subject of the scene description.
- Objects that describe the details of the scene: who/whom, where, when, what, how, etc.
- Relationships between scenes: then, if, because, etc.
- time: an absolute time when the scene occurred (xsd:DateTime).
- source: original text of the scene (English/Japanese; literal).

3.2. Procedure of knowledge graph construction

The following procedure was used to convert the eight mystery stories into knowledge graphs:

1. Extract sentences necessary for deduction from mystery stories (in Japanese) whose copyrights have expired. For each novel, about 300 to 500 sentences were extracted manually. Those sentences were selected mainly where they were closely related to the identification of the culprits.

[3] Only scenes that are judged to be necessary for the deduction are converted to knowledge graphs.

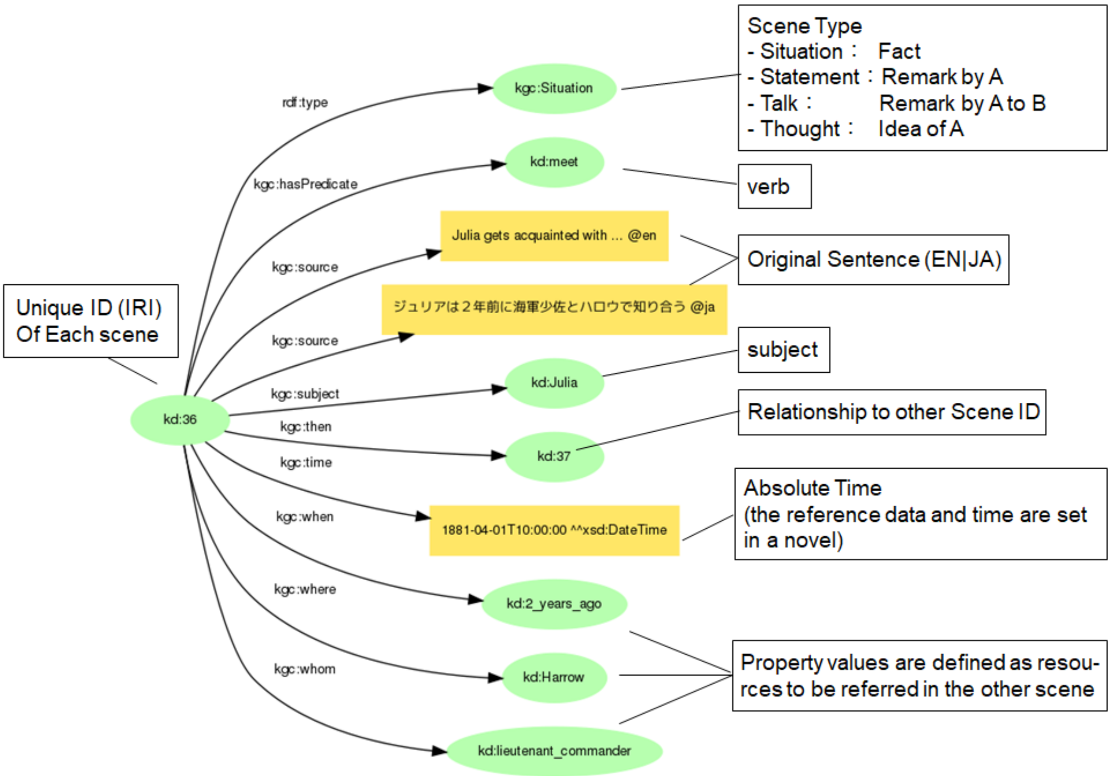

Figure 2: Example of a scene in the knowledge graph of mystery stories.

2. Rewriting the original text into sentences with clear a subject and object (i.e., short sentences). One short sentence corresponds to one scene on the knowledge graph.
3. Assign semantic roles (e.g., 5W1H) to phrases using natural language processing tools. The results are output as a predicate and an object for each scene in a spreadsheet, and are visually checked at the end.
4. Control orthographical variants. We eliminate any notational distortions on a novel-by-novel basis and across novels as much as possible during the construction phase.
5. Add relationships between scenes (e.g., temporal relationships).
6. Translate the source text into English and convert the entire text into a knowledge graph.

Note that the series of tasks were performed by part time students and software engineers (general programmers, not advanced knowledge engineers). The costs for knowledge graph construction of each story are as follows: (Step 1) 3 hours per person, (Step 2) 20 hours per person, (Step 3) 5 hours per person, (Step 4) 7 hours per person, (Step 5) 3 hours per person, and (Step 6) about 1 hour per person.

In addition to the above procedure, we have made a series of improvements to the content of the descriptions in the knowledge graph. In [5] , we present a guideline consisting of ten items/steps we found through the refinement process of the knowledge graphs.

3.3. Example queries

The latest version of the knowledge graph constructed for the Challenges is available on GitHub[4] and provides a SPARQL endpoint for searching. Here, we show some examples of SPARQL queries for the knowledge graph focusing on scenes. In the SPARQL endpoint, graph IRI is set for each novel. Therefore, the users can specify the target novels for searching using FROM clause.

Listing1 shows a SPARQL query for obtaining all triples of scene 036 in “The Speckled Band”. Because each scene is represented by triples whose subjects are IRI of a scene, the users can obtain details of scene descriptions to get its property-object pairs. If a kind of property is specified in the query (e.g., kgc:hasPredicate), its object is obtained.

Contrary to Listing1, by specifying a property-object pair, it is possible to retrieve scenes containing that combination. For example, Listing2 shows a SPARQL query for obtaining all scenes whose subjects (performers of the action in the scene) are Holmes in “The Speckled Band”.

Listing 1: Get all triples of Scene 036 in “The Speckled Band”

SPARQL query:

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX kgc: <http://kgc.knowledge-graph.jp/ontology/kgc.owl#>
PREFIX pred: <http://kgc.knowledge-graph.jp/data/predicate/>
PREFIX sb: <http://kgc.knowledge-graph.jp/data/SpeckledBand/>

SELECT ?p ?o
FROM <http://kgc.knowledge-graph.jp/data/SpeckledBand>
WHERE {
  sb:036 ?p ?o .
}
```

Query result:

?s	?p
kgc:source	”ジュリアは２年前に海軍少佐とハロウで知り合う”@ja
kgc:source	”Julia gets acquainted with Major Navy two years ago at Harrow”@en
rdf:type	kgc:Situation
kgc:hasPredicate	pred:meet
kgc:subject	sb:Julia
kgc:when	sb:2_years_ago
kgc:when	sb:1880-12-24T10
kgc:where	sb:Harrow
kgc:whom	sb:lieutenant_commander
kgc:time	”1880-12-24T10:00:00”^^xsd:dateTime

[4]https://github.com/KnowledgeGraphJapan/KGRC-RDF/tree/ikgrc2023

Listing 2: Get all scenes whose subjects (performers of the action in the scene) is Holmes in "The Speckled Band"

SPARQL query:

```
PREFIX kgc: <http://kgc.knowledge-graph.jp/ontology/kgc.owl#>
PREFIX sb: <http://kgc.knowledge-graph.jp/data/SpeckledBand/>

SELECT ?s
FROM <http://kgc.knowledge-graph.jp/data/SpeckledBand>
WHERE {
  ?s kgc:subject sb:Holmes .
}
```

Query result(parts):

?s
sb:131
sb:132
sb:140
sb:142
sb:143
sb:150
...

Listing3, shows a query to obtain all relationships among scenes in "The Speckled Band". In this query, subjects and objects of triples are specified as scenes (sub-class of kgc:Scene). These relationships represent the flow of the story. It is an important feature of this knowledge graph of mystery stories.

Listing 3: Get all relationships among scenes in "The Speckled Band".

SPARQL query:

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX kgc: <http://kgc.knowledge-graph.jp/ontology/kgc.owl#>
PREFIX sb: <http://kgc.knowledge-graph.jp/data/SpeckledBand/>

SELECT ?s ?p ?o
FROM <http://kgc.knowledge-graph.jp/data/SpeckledBand>
WHERE {
  ?s rdf:type/rdfs:subClassOf kgc:Scene .
  ?o rdf:type/rdfs:subClassOf kgc:Scene .
  ?s ?p ?o .
}
```

Query result(parts):

?s	?p	?o
sb:001	kgc:at_the_same_time	sb:002
sb:002	kgc:at_the_same_time	sb:003
sb:014	kgc:then	sb:015

sb:015	kgc:then	sb:016
sb:016	kgc:if	sb:017

...

Listing4 is an example to search across different stories. Because the same action is represented using the same IRI across all stories, it is possible to search across them by specifying the IRI of action. Its targets are selected using FROM clause.

Listing 4: Get all scenes whose predicate is "meet" across the eight stories with its subjects and target objects.

SPARQL query:

```
PREFIX kgc: <http://kgc.knowledge-graph.jp/ontology/kgc.owl#>
PREFIX pred: <http://kgc.knowledge-graph.jp/data/predicate/>
PREFIX sb: <http://kgc.knowledge-graph.jp/data/SpeckledBand/>
PREFIX df: <http://kgc.knowledge-graph.jp/data/DevilsFoot/>
PREFIX dm: <http://kgc.knowledge-graph.jp/data/DancingMen/>
PREFIX ci: <http://kgc.knowledge-graph.jp/data/ACaseOfIdentity/>
PREFIX cm: <http://kgc.knowledge-graph.jp/data/CrookedMan/>
PREFIX ag: <http://kgc.knowledge-graph.jp/data/AbbeyGrange/>
PREFIX rp: <http://kgc.knowledge-graph.jp/data/ResidentPatient/>

SELECT DISTINCT ?s ?subj ?obj
FROM <http://kgc.knowledge-graph.jp/data/SpeckledBand>
FROM <http://kgc.knowledge-graph.jp/data/DevilsFoot>
FROM <http://kgc.knowledge-graph.jp/data/SilverBlaze>
FROM <http://kgc.knowledge-graph.jp/data/DancingMen>
FROM <http://kgc.knowledge-graph.jp/data/ACaseOfIdentity>
FROM <http://kgc.knowledge-graph.jp/data/CrookedMan>
FROM <http://kgc.knowledge-graph.jp/data/AbbeyGrange>
FROM <http://kgc.knowledge-graph.jp/data/ResidentPatient>
WHERE {
  ?s kgc:hasPredicate pred:meet .
  ?s kgc:subject ?subj.
  ?s kgc:whom ?obj.
}
```

Query result(parts):

?s	?subj	?obj
cm:077	cm:Barclay	cm:Nancy
cm:178	cm:Nancy	cm:Henry
cm:178	cm:Morrison	cm:Henry
cm:240	cm:Holmes	cm:Henry
cm:240	cm:Watson	cm:Henry
sb:036	sb:Julia	sb:lieutenant_commander
sb:081	sb:Helen	sb:Roylott
sb:152	sb:Holmes	sb:Helen
sb:152	sb:Watson	sb:Helen
df:438	df:Sterndale	df:Mortimer

Table 1
Comparisons of main properties of scenes across target novels.

Novel	number of scenes	hasPredicate	hasProperty	hasPart	subject	infoSource	infoReceiver
The Speckled Band	401	348	53	9	435	40	1
The Dancing Men	212	179	31	0	217	47	0
A Case of Identity	544	449	86	0	464	396	0
The Devil's Foot	480	457	23	0	559	320	0
The Crooked Man	354	300	55	0	402	152	0
The Abbey Grange	416	365	67	0	475	316	0
Silver Blaze	394	353	42	0	425	0	0
The Resident Patient	324	209	115	0	376	0	0
total	3125	2660	472	9	3353	1271	1

df:438	df:Sterndale	df:George
df:438	df:Sterndale	df:Owen
df:455	df:Sterndale	df:Holmes
ci:134	ci:Sutherland	ci:Hosmer
ci:145	ci:Sutherland	ci:Hosmer
ci:147	ci:Sutherland	ci:Hosmer
ci:389a	ci:Holmes	ci:Windibank

3.4. Published Knowledge Graphs

Knowledge graphs of mystery stories were constructed based on the above schema and published as open datasets for the Knowledge Graph Reasoning Challenge. They were extended and refined through five challenges according to feedbacks from participants and discussion among organizers. Some of the findings from these processes have been compiled into guidelines for knowledge graph construction [5].

Here we compare the published knowledge graphs of the eight mystery stories and describe their appearance. Since these knowledge graphs are structured around scenes, comparisons between novels are also made around scenes. First, we compare the properties used to define scenes. Then, the relationships between scenes are compared, and finally the predicates (actions) used in each scene are compared.

Table 1 compares number of scenes in each novel and main properties used to define scenes. These knowledge graphs consist of around 300 to 500 scenes while "The Dancing Men" has only 200 scenes. It is because the main topic of "The Dancing Men" is deciphering the cipher. Each scene must have *hasPredicate* or *hasProperty* property. *hasPredicate* represent a predicate that describes the content of the scene and *hasProperty* represents a property of the person or thing that is the subject of the scene description. Subjects of the predicate/property are described by *subject* property. The number of these three properties does not vary much among novels. *infoSource* property is used to describe information sources when its scene type is Thought or Statement. *infoReceiver* property shows receiver of some remarks while it is used only once across all novels. *hasPart* property shows parts of some scenes while it is used only in "The Speckled Band".

Table 2
Comparisons of 5W1H properties of scenes across target stories.

Novel	how	what	when	whom	why	time
The Speckled Band	45	186	110	30	32	95
The Dancing Men	11	151	100	22	2	81
A Case of Identity	23	769	60	54	9	8
The Devil's Foot	19	541	76	27	11	3
The Crooked Man	22	312	75	59	9	39
The Abbey Grange	39	256	76	0	38	22
Silver Blaze	30	500	90	0	54	26
The Resident Patient	20	404	46	0	2	18
total	209	3119	633	192	157	292

Table 3
Comparisons of location properties of scenes across target stories.

Novel	where	from	to	right	left	middle	on	near	next_to	opposite	adjunct
The Speckled Band	97	34	55	2	2	1	7	3	6	2	3
The Dancing Men	34	16	14	1	1	0	9	0	0	1	0
A Case of Identity	50	15	20	0	0	0	8	0	0	0	0
The Devil's Foot	49	21	45	0	0	0	16	4	0	1	0
The Crooked Man	51	11	19	0	0	0	1	1	0	0	0
The Abbey Grange	65	15	41	1	1	0	12	8	0	1	0
Silver Blaze	41	15	81	0	0	0	12	12	0	0	0
The Resident Patient	21	7	45	0	0	0	6	0	0	0	0
total	408	134	320	4	4	1	71	28	6	5	3

Table 4
Comparisons of relationships between scenes.

Novel	at_the_ same_time	if	because	then	therefore	how	what	when	where	why	subject	infoSource
The Speckled Band	9	2	0	124	4	4	10	7	0	21	0	0
The Dancing Men	4	0	12	14	0	0	1	1	0	0	0	0
A Case of Identity	3	0	0	11	0	0	9	2	0	0	0	0
The Devil's Foot	0	1	0	14	1	0	36	0	0	0	0	0
The Crooked Man	1	1	0	3	0	0	31	0	0	1	4	13
The Abbey Grange	0	1	5	19	0	0	1	5	0	3	0	0
Silver Blaze	22	0	2	173	0	2	275	28	1	42	0	0
The Resident Patient	41	4	21	116	2	0	264	2	0	0	1	0
total	80	9	40	474	7	6	627	45	1	67	5	13

Table 2 compares properties to represent 5W1H of scenes. In these properties, *what* is used most frequently, followed by *why*. On the other hand, *whom* is not used at all in the three novels. In this knowledge graph, *time* is a property that represents absolute time, but its frequency of use varies from novel to novel. This may be due to differences in the representation of time in the stories.

Table 3 compares properties to represent location of scenes. In these properties, *where* is used most frequently, followed by *to* and *from*. This is probably because these three properties are used to represent basic location while there are properties to represent more detailed location.

Table 4 compares properties to represent relationships between scenes. The use of *what* is extremely frequent in the two novels. This may be because *what* is often used to divide a long

Table 5
Comparison of the top 10 most frequently used predicates across the eight novels.

Novel	say	have	exist	go	think	see	hear	equalTo	enter	meet
The Speckled Band	6	28	40	14	0	8	14	0	1	3
The Dancing Men	11	10	9	3	2	5	1	0	1	0
A Case of Identity	73	20	9	11	11	5	2	2	2	7
The Devil's Foot	31	19	19	11	11	7	7	2	10	3
The Crooked Man	15	16	10	10	13	8	7	20	5	4
The Abbey Grange	15	13	18	6	4	7	5	1	2	9
Silver Blaze	31	20	11	6	0	2	2	0	0	3
The Resident Patient	20	5	1	3	2	0	1	6	9	0
Total	202	131	117	64	43	42	39	31	30	29

sentence into multiple scenes, but the usage of the two novels needs to be closely examined. Other than *what*, *then* is the most commonly used to indicate the order of scenes. This is not surprising, since the order of scenes is important in a story. The use of *because* and *why* to express reasons varies widely from novel to novel.

Finally, we compare the use of predicates for actions. In the eight novels as a whole, 665 different actions are used. Of these actions, 53 were used more than 10 times. Table 5 compares the top 10 most frequently used predicates across the eight novels. The table shows that predicates for basic actions such as *say*, *have*, and *exist* are frequently used. However, there are differences in the actions used in each novel.

3.5. Example Works using the Knowledge Graphs

Through the challenges, various approaches have been proposed by researchers from private companies and universities to develop AI systems that can infer and explain the truth of mysterious crimes. The lists of proposed works and summaries of each work are shown in the [5]. There are two major approaches to use the knowledge graphs for reasoning. One is logical explanation based on ontology and the other is interpretation of knowledge graph embedding.

For example, Ugai et al. (IKGRC2023, Main Track) [3, 6] considered that the analysis of motive, opportunity, and means is necessary to identify the culprit, and they prepared this knowledge as external knowledge and combined it with the event-centric knowledge graphs provided in this challenge to make inferences. Specifically, ontologies about the motive and means of murder are manually created and used for inference. The knowledge graphs were extended with external knowledge by linking these ontologies to the knowledge graphs of mystery stories via ConceptNet and WordNet. The extended knowledge graphs were embedded in a low-dimensional vector space to estimate the scores that each character could possess for the motive and means of the crime. These scores were then utilized to infer the culprit. Ontologies of motives and means help logically explain the reasons deduced.

Kurokawa (KGRC2020, Main Track) [7] converted the event-centric knowledge graphs provided in this challenge into the standard triples and proposed an approach to predict the culprits of the stories using link prediction with multiple knowledge graph embedding methods

(TransE [8], PTransE [9], R-GCN [10]). To integrate multiple KGs, Kurokawa applied parameter sharing proposed in ITransE, which shares the same embeddings for common entities, e.g., ”Holmes” and ”Watson,” and for common relations, e.g., ”kill.” In addition, an external commonsense KG, ConceptNet [11], was introduced to enrich semantic information. Kurokawa also conducted an experiment using XKE [12] and GNNExplainer [13] to show plausible graph paths as a basis for the prediction results.

4. Related Work

Knowledge graphs can be used to describe static relationships between things, such as in product data, a thesaurus, and human relationships, as well as events that occur in space and time, such as observational data. In recent years, knowledge graphs of events or scenes, such as video content [14], historical fact [15], news [16], and narrative content [17] have been actively studied. VirtualHome2KG [14] can generate event-centric knowledge graphs of the content of videos of daily activities simulated in virtual space and demonstrate various applications, including accident risk detection. EventKG [15] is a knowledge graph describing 690,000 contemporary and historical events and incidents for the purpose of answering questions and generating histories (timelines) from a specific perspective. The schema is based on the aforementioned Simple Event Model (SEM) [18] and is extended to express temporal relationships, and so forth. It has many similarities with our schema, such as definitions of relationships among events. However, the granularity of its target events is considerably larger than in our scenes, and it is difficult to represent information such as who, when, and how for each scene using EventKG’s model (although it is possible to describe them, it would be a complex graph, and it would be difficult to construct and search the dataset). ECKG [16] provides its own model to annotate information extracted when building a knowledge graph directly from news events written in a natural language. It provides a unique model. However, the model is simple (only who, what, where, and when) because automatic extraction is the subject matter. Drammer [17] is not simply a chronological representation and comparison of narrative content, but is fiction–specific. It is an ontology that includes conflicts between characters, segmentations of the narrative, and definitions of emotion and belief for more dramatic representations. It was constructed by analyzing many dramas, but its purpose is different from that of this study, which is intended to represent facts (including falsehoods) in the real world.

By contrast, in this study we (1) constructed knowledge graphs that convert the background of the case and the characters into knowledge, using a mystery novel as a subject, and (2) conducted a technical challenge to identify correctly the culprit and cause of a case or an accident from given information using inference and estimation techniques to explain the reasons (evidence, tricks, etc.) for such identification appropriately.

As for this technical challenge, in top conferences of AI and neural networks, such as IJCAI, AAAI, NIPS, and ICML, papers and workshops that have “explainability” as a keyword and that analyze the properties of AI models have significantly increased since 2016. However, no other research activity exists like the challenge discussed in this work, which uses knowledge graphs, including social problems as common test sets, and tries to solve the problems with explainability (i.e., using XAI), aiming to integrate inductive estimation and deductive reasoning.

5. Conclusion

In this paper, we discussed datasets of mystery stories for the Knowledge Graph Reasoning Challenge. Its feature is a knowledge graph schema focusing on scenes of novels so that the user can search its contents according to the flow of the stories. They were built on eight Sherlock Holmes short stories and published as open datasets.

Through comparison of the knowledge graphs across novels, we discussed how the proposed schema and vocabularies are used in them. Each novel has a narrative flow that makes sense to the reader, and these are structured as a knowledge graph. Therefore, we believe that such a comparison offers suggestions for considering the representation of stories as knowledge. In this comparison, we observed that there is variation in properties and the vocabulary of actions from novel to novel. Based on the results of this comparison, it is necessary to review the schema, systematize the vocabulary, and examine the policy for constructing a knowledge graph.

Acknowledgments

This paper is based on results obtained from a project, JPNP20006, commissioned by the New Energy and Industrial Technology Development Organization (NEDO), and Japan Society Promotion of Science (JSPS) KAKENHI Grant Numbers JP19H04168 and JP22K18008.

References

[1] 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, K. Kozaki, Report on the First Knowledge Graph Reasoning Challenge 2018 - Toward the eXplainable AI System -, 9th Joint Int'l Semantic Tech. Conf. (JIST2019) (2019) 18–34.

[2] K. Kozaki, S. Egami, K. Matsushita, T. Ugai, T. Kawamura, Knowledge graph reasoning techniques through studies on mystery stories -report on the knowledge graph reasoning challenge 2018 to 2020, 1st Int'l Workshop on Knowledge Graph Reasoning for Explainable Artificial Intelligence (KGR4XAI) (2021).

[3] T. Ugai, Y. Koyanagi, F. Nishino, A logical approach to criminal case investigation, 1st Int'l Workshop on Knowledge Graph Reasoning for Explainable Artificial Intelligence (KGR4XAI) (2021).

[4] S. Katsushima, H. Anada, S. Egami, K. Fukuda, A criminal detection of mystery novel using the principal components regression analysis considering co-occurrence words, 1st Int'l Workshop on Knowledge Graph Reasoning for Explainable Artificial Intelligence (KGR4XAI) (2021).

[5] T. Kawamura, S. Egami, K. Matsushita, T. Ugai, K. Fukuda, K. Kozaki, Contextualized scene knowledge graphs for xai benchmarking, IJCKG '22, Association for Computing Machinery, New York, NY, USA, 2023, pp. 64–72. URL: https://doi.org/10.1145/3579051.3579061. doi:10.1145/3579051.3579061.

[6] T. Ugai, Criminal investigation with augmented ontology and link prediction, in: Proceedings of the 1st International Knowledge Graph Reasoning Challenge co-located with 17th IEEE International Conference on Semantic Computing, 2023. To appear.
[7] M. Kurokawa, Explainable knowledge reasoning framework using multiple knowledge graph embedding, in: Proceedings of the 10th International Joint Conference on Knowledge Graphs, IJCKG '21, Association for Computing Machinery, New York, NY, USA, 2022, pp. 172–176. doi:10.1145/3502223.3502248.
[8] A. Bordes, N. Usunier, A. Garcia-Duran, J. Weston, O. Yakhnenko, Translating embeddings for modeling multi-relational data, in: C. Burges, L. Bottou, M. Welling, Z. Ghahramani, K. Weinberger (Eds.), Advances in Neural Information Processing Systems, volume 26, Curran Associates, Inc., 2013. URL: https://proceedings.neurips.cc/paper/2013/file/1cecc7a77928ca8133fa24680a88d2f9-Paper.pdf.
[9] Y. Lin, Z. Liu, H. Luan, M. Sun, S. Rao, S. Liu, Modeling relation paths for representation learning of knowledge bases, in: Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, Lisbon, Portugal, 2015, pp. 705–714. URL: https://aclanthology.org/D15-1082. doi:10.18653/v1/D15-1082.
[10] M. Schlichtkrull, T. N. Kipf, P. Bloem, R. vanÂ den Berg, I. Titov, M. Welling, Modeling relational data with graph convolutional networks, in: A. Gangemi, R. Navigli, M.-E. Vidal, P. Hitzler, R. Troncy, L. Hollink, A. Tordai, M. Alam (Eds.), The Semantic Web, Springer International Publishing, Cham, 2018, pp. 593–607.
[11] R. Speer, J. Chin, C. Havasi, ConceptNet 5.5: An open multilingual graph of general knowledge, in: Proceedings of the AAAI conference on artificial intelligence, volume 31, 2017.
[12] A. C. Gusmão, A. H. C. Correia, G. De Bona, F. G. Cozman, Interpreting embedding models of knowledge bases: a pedagogical approach, arXiv preprint arXiv:1806.09504 (2018).
[13] Z. Ying, D. Bourgeois, J. You, M. Zitnik, J. Leskovec, Gnnexplainer: Generating explanations for graph neural networks, Advances in neural information processing systems 32 (2019).
[14] S. Egami, T. Ugai, M. Oono, K. Kitamura, K. Fukuda, Synthesizing event-centric knowledge graphs of daily activities using virtual space, IEEE Access 11 (2023) 23857–23873. doi:10.1109/ACCESS.2023.3253807.
[15] S. Gottschalk, E. Demidova, EventKG - the Hub of Event Knowledge on the Web - and Biographical Timeline Generation, Semantic Web J. 10 (2019) 1039–1070.
[16] M. Rospocher, M. van Erp, P. Vossen, A. Fokkens, I. Aldabe, G. Rigau, A. Soroa, T. Ploeger, T. Bogaard, Building event-centric knowledge graphs from news, J. of Web Semantics 37-38 (2016) 132–151. doi:10.1016/j.websem.2015.12.004.
[17] V. Lombardo, R. Damiano, A. Pizzo, Drammar: A comprehensive ontological resource on drama, 17th Int'l Semantic Web Conf. (ISWC2018) (2018) 103–118. doi:10.1007/978-3-030-00668-6_7.
[18] W. van Hage, D. Ceolin, The simple event model, Situation Awareness with Systems of Systems (2012) 149–169. doi:10.1007/978-1-4614-6230-9_10.