

# ナレッジグラフ推論チャレンジ 2023 応募シート

## 1. 応募者に関する情報

- 氏名またはチーム名：青山 仁
- 所属：青山学院大学 森田研究室
- メールアドレス（代表）：[c5623213@aoyama.jp](mailto:c5623213@aoyama.jp)
  - 応募者に学生が含まれる：はい
  - 応募者の代表が学生である：はい

## 2. 応募部門：一般部門

## 3. 構築したナレッジグラフについて

### ● 構築したナレッジグラフ

- テーマ：日常生活行動ナレッジグラフ
- 説明：家庭内のエージェントが「いつ」、「どこで」、「誰が」、「どのような対象物に対して」、「どのような行動を」行ったか、そしてその結果としての「状態」や「位置」を表すナレッジグラフを構築した。

### ● 構築したナレッジグラフの基本情報

- 構築したナレッジグラフの数：12
- データサイズ：10,169~39,527 トリプル, 533~1,209 KB
- データ形式：RDF (Turtle)

### ● 構築したナレッジグラフの入手先

- 公開先 URL：<https://github.com/JinAoyama/kgrc2023>

#### 4. ナレッジグラフ構築に用いた「言語モデル」および「構築手法」について

- ナレッジグラフ構築に用いた「言語モデル」
  - LLM : GPT-3.5 (gpt-3.5-turbo-0613)
  - URL : <https://platform.openai.com/docs/models/gpt-3-5>
- ナレッジグラフ構築に用いた「データ」
  - VirtualHome (VH) データセット [\[Puig+ 2018\]](#)
    - データの説明 : データの詳細や利用方法は構築手法で説明する.
    - データの形式 : テキストファイルに「アクティビティ名」, 「行動の説明文」, 「アクションスクリプト」が記載される.

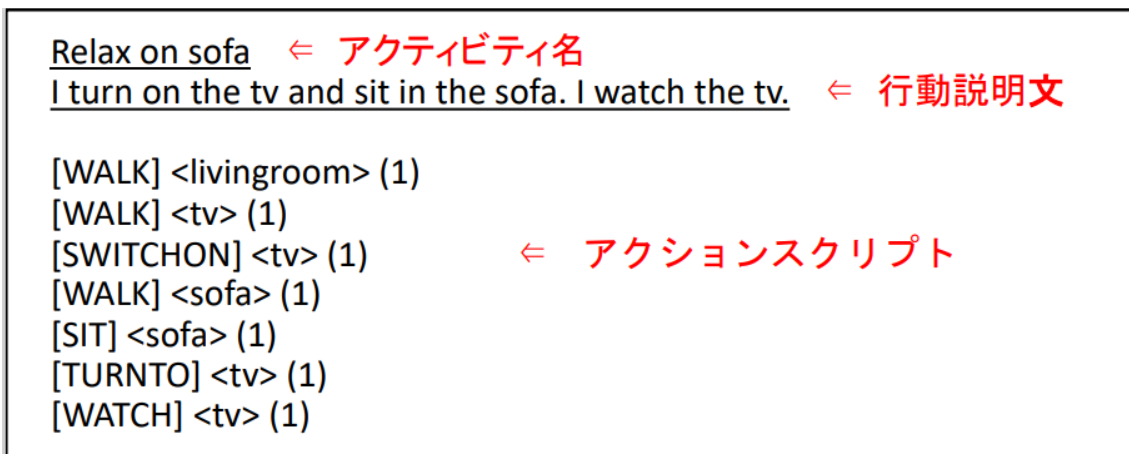


図 1. VirtualHome データセットのサンプル

- ナレッジグラフの構築手法の説明
  - ※ ナレッジグラフの構築には VirtualHome2KG [\[Egami+ 2023\]](#) を利用する.
  - VirtualHome2KG を利用したナレッジグラフの構築
    - VirtualHome2KG は, 家庭内の日常生活行動を仮想空間で実行できる VirtualHome (VH) を利用し, “ソファでくつろぐ”, “テレビを見る”などの日常生活における行動 (以降, アクティビティ) のシミュレーション情報をナレッジグラフとして合成するシステムである.
    - VirtualHome2KG を利用したナレッジグラフの構築には, VH 上でシミュレーションを実行して, 行動の履歴を蓄積する必要がある.

- VH の実行には、VH データセットの一部であるアクションスクリプト（アクション、オブジェクト名、オブジェクトの ID の組のシーケンスとして表現）を用いる。つまり、VH データセットと VirtualHome2KG を利用することでナレッジグラフの構築ができる。
- 今回提案するのは、直接的にナレッジグラフを構築する手法ではなく、アクティビティを入力として、アクションスクリプトを生成する手法である。また、提案手法により構築したデータセットと VirtualHome2KG を利用してナレッジグラフを構築する。

#### - アクションスクリプトの生成（データセットの構築）

##### - 概要：

提案システムの構成図を図 2 に示す。図 2 の Comp.1 は、LLM を用いて入力となるアクティビティの詳細な行動説明文を生成する。Comp.2 は、Comp.1 が生成した詳細な行動説明文をもとに、LLM を用いてアクションのステップを生成する。Comp.3 は、Comp.2 が生成したアクションのステップを SentenceTransformers の文書間類似度の技術を用いてアクションスクリプトに変換する。Comp.4 は生成したアクションスクリプトの実行の検証と必要に応じて修正を行う。Comp.4 はアクションスクリプトの修正をする必要がなくなる（アクションスクリプトが実行可能になる）まで繰り返す。

※ 提案手法は過去発表したもの[\[青山+ 2023\]](#)であるため、より詳細な内容は[\[青山+ 2023\]](#)を参照。

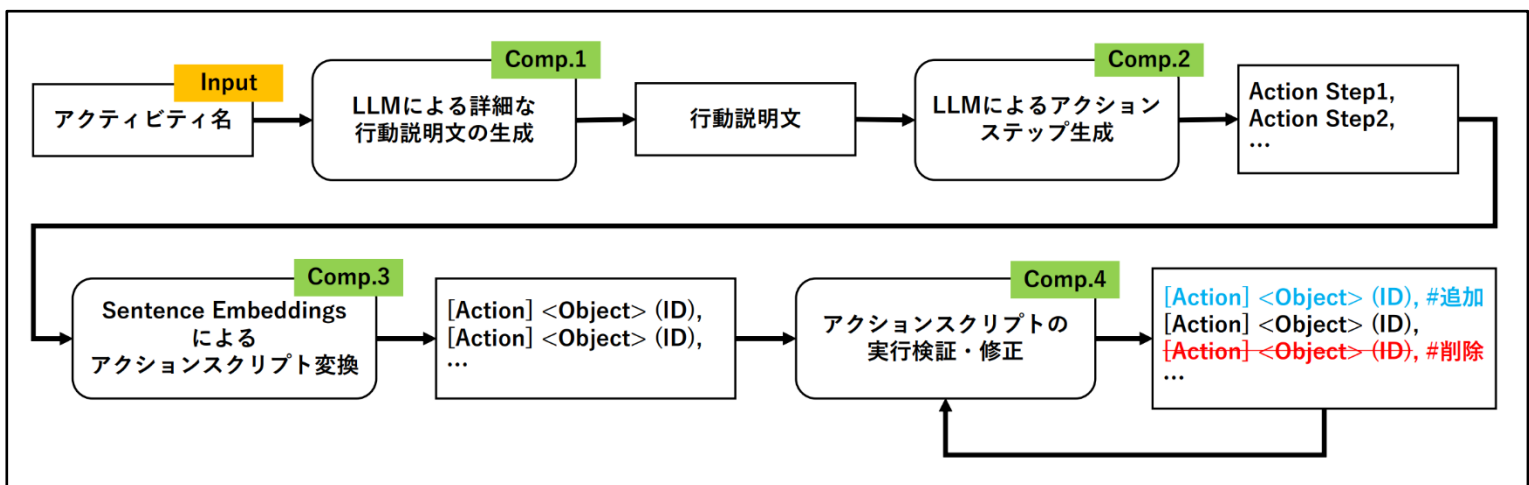


図 2. 提案システム構成図

```

(“system”, “Describe the details of the activity input, while keeping to the following restrictions
1. You can use only the following actions in the description:
Find, Walk, Run, ...

2. You can use only the following objects and rooms in the description:
There are the following objects in the bathroom:
barsoap, bathroomcabinet, ...

There are the following objects in the bedroom:
bed, book, ...

... “),
(“human”, “Describe the details of the following activity: {activity} ”)

```

図 3. アクティビティの詳細な行動説明文を生成するプロンプト (Comp.1)

### Comp.1. LLM による詳細な行動説明文の生成

- Zero-shot 学習, Few-shot 学習, Fine-tuning の 3 つの方法を提案する. 学習データには VH データセットを用いる.

- **Zero-shot 学習 :**

図 3 のプロンプトを用いて LLM に例を与えずに, 入力となるアクティビティの詳細な行動説明文を生成する.

具体的には VH で実行可能な行動と VH の環境内に存在するすべての物や部屋を提示して, LLM の出力に制限をかけて生成させている. また, {activity} は変数であり, 入力となるアクティビティを代入する.

- **Few-shot 学習 :**

LLM に例を与えて, 入力となるアクティビティの詳細な行動説明文を生成する.

具体的には, 図 3 のプロンプトに追加で入出力の例を記述したプロンプト (図 4) を構築して与える. 図 4 のプロンプトの {activity}, {description} は変数であり, アクティビティと, アクティビティの詳細な行動説明文を代入する. 例の与え方として 2 パターンの方法を提案する.

```

(“human”, “Describe the details of the following activity: {activity} ”),
(“ai”, “{description} ”

```

図 4. Few-shot 学習で与える例のプロンプトのテンプレート (Comp.1)

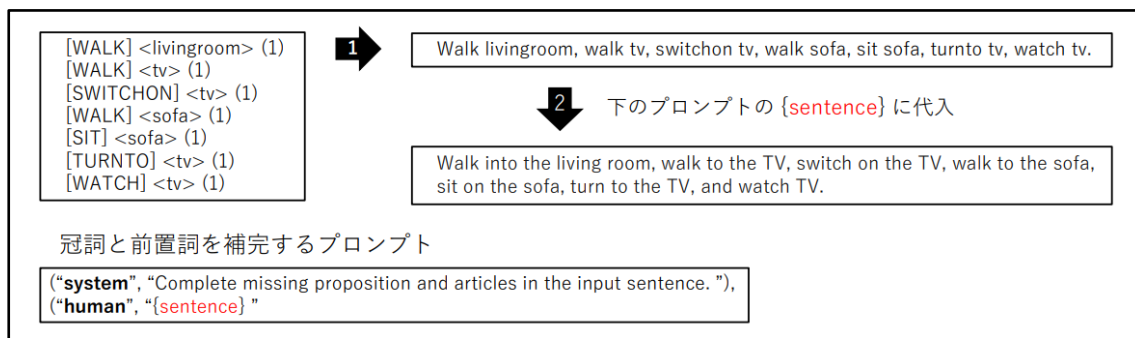


図 5. アクションスクリプトを説明文に変換するプロセス (Comp.1)

1 つ目は、VH データセットの 1 行目のアクティビティ名と、2 行目の行動説明文を例にする方法。

2 つ目は、VH データセットの 1 行目のアクティビティ名と、4 行目以降のアクションスクリプトを文章に変換したものを行動説明文として例にする方法である。アクションスクリプトを文章に変換する方法は図 5 の通りであり、まず、アクションスクリプトに含まれる ID やカッコを除き一つの文にする。その後、図 5 の冠詞と前置詞を補完するプロンプトを用いて冠詞と前置詞を補完した文章を LLM に生成させる。

#### - Fine-tuning :

Few-shot 学習と同様に 2 パターンで例を構築して、構築した例（プロンプト）を JSONAL 形式に変換して Fine-tuning する。

#### Comp.2. LLM によるアクションステップ生成

- Comp.1 で生成した行動説明文を、図 6 のプロンプトを用いて LLM にアクションのステップを生成させ、リスト形式に変換する。また、図 6 のプロンプトの{**description**}は変数であり、行動説明文を代入する。

#### - Comp.2 の入出力の例 :

入力 : “pick up phone and sit on the sofa”

出力 : 1. Grab cellphone. 2. Walk to sofa. 3. Sit on sofa.

→[“Grab cellphone”, “Walk to sofa”, “Sit on sofa”]

(“**system**”, “Each action step should be briefly described while keeping to the following restrictions.  
 1. You can use only the following actions in the description:  
 Find, Walk, Run, ...  
 2. You can use only the following objects and rooms in the description:  
 There are the following objects in the bathroom:  
 barsoap, bathroomcabinet, ...  
 There are the following objects in the bedroom:  
 bed, book, ...  
 ... “),  
 (“**human**”, “Rewrite the following description as action steps: {description} ”)

図 6. アクションのステップを生成するプロンプト (Comp.2)

### Comp.3. Sentence Embeddings によるアクションスクリプト変換

- **Hash Table (表 1) の作成 :**  
 VH 上で実行可能な行動と、環境内に存在するすべての物や部屋から成り得る行動文を Key として作成する。  
 作成した行動文に対応するアクションスクリプトを Value として作成する。
- **アクションスクリプト変換**  
 Comp.2 で生成したアクションのステップに対して、作成した辞書 (表 1) 内で最も類似度の高い行動文 (Key) を選択して、それに対応するアクションスクリプト (Value) に変換する。
- **Comp.3 の入出力の例**  
 入力 : [“Grab cellphone”, “Walk to sofa”, “Sit on sofa”]  
 出力 : [“[grab] <cellphone> (1)”, “[walk] <sofa> (1)”, “[sit] <sofa> (1)”]

Key	Value
Walk bathroom	[walk] <bathroom> (1)
Walk bedroom	[walk] <bedroom> (1)
...	...
Sit chair	[sit] <chair> (1)
Sit sofa	[sit] <sofa> (1)
...	...

表 1. アクションスクリプト変換のための Hash Table ※ 一部省略

#### Comp.4. アクションスクリプトの実行検証・修正

- アクションスクリプトの実行検証

シミュレータを用いて、実行可能か検証し、実行できなかった場合は、失敗したアクションステップを特定しエラーメッセージを出力させる。

- アクションスクリプトの修正

エラーメッセージが、アクションの対象となるオブジェクトとの距離が遠いことが原因で実行できないことを示唆する場合、対象オブジェクトにアプローチするアクション（"Walk"）を追加して修正する。

その他の原因の場合には、失敗したアクションのステップを削除する。

- Comp.4 の入出力の例

入力：["[grab] <cellphone> (1)", "[walk] <sofa> (1)",  
"[sit] <sofa> (1)"]

出力：

["[walk] <cellphone> (1)", ⇐ "cellphone" に近づく動作追加

"[grab] <cellphone> (1)", ⇐ "cellphone" と距離が遠くて実行不可

["[walk] <sofa> (1)",

"[sit] <sofa> (1)"]

- パフォーマンス情報

- 使用した PC のスペック

- プロセッサ (CPU) : 3.60 GHz

- 実装 RAM (メモリ) : 64.0 GHz

- 参考情報

- [Puig+ 2018] Xavier Puig, et al. Virtualhome: Simulating household activities via programs. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 8494–8502, 2018.
- [Egami+ 2023] Shusaku Egami, et al. Synthesizing event-centric knowledge graphs of daily activities using virtual space. *IEEE Access*, Vol. 11, pp. 23857–23873, 2023.
- VirtualHome2KG
  - GitHub URL : <https://github.com/aistairc/VirtualHome2KG>
  - 利用方法は, GitHub の README 参照
- [青山+ 2023] 青山仁, 森田 武史, 鵜飼 孝典, 江上 周作, 福田 賢一郎 (2023) 「LLM を活用し抽象的なタスク記述からの VirtualHome のためのアクションスクリプト自動生成」, 第 61 回人工知能学会セマンティックウェブとオントロジー研究会, SIG-SWO-061-10, 1-10  
DOI : [https://doi.org/10.11517/jsaisigtwo.2023.SWO-061\\_10](https://doi.org/10.11517/jsaisigtwo.2023.SWO-061_10)



## 5. 構築したナレッジグラフの評価

- 入力となるアクティビティをナレッジグラフで表現できているかを評価する必要があるが、VirtualHome2KG を利用して構築するナレッジグラフは、利用するデータに依存するため、主に提案システムの評価を行う。

### - 提案システムの評価

以下の3つを VH データセットとシミュレータを用いて評価する。

- アクティビティから生成した詳細な行動説明文の正しさ
- アクティビティから生成したアクションスクリプトの正しさ
- アクティビティから生成したアクションスクリプトの実行率

- **評価環境**：VH はシーン 1 から 7 の 7 種類の提供されている家屋のうちのシーン 1 を対象とする。

- **評価データ**：VH データセットを利用する。（データ数：197 個）

### - 評価指標

1. **ROUGE**：N-gram 単位で一致度を取る手法（uni-gram 単位とする）

2. **LCSscore**：LCS(Longest Common Subsequence)をもとに次のように定義

$$LCSscore = \frac{LCS(Generated\ Action\ Script, \ Correct\ Action\ Script)}{\max(length(Generated\ Action\ Script), \ length(Correct\ Action\ Script))}$$

3. **Correctness**：行動の順序は考慮せず一致したアクションステップの数（正解数）をとり、以下の式でスコアを計算する。

$$Precision = \frac{Number\ of\ Correct\ Action\ Steps}{length(Generated\ Action\ Script)}$$

$$Recall = \frac{Number\ of\ Correct\ Action\ Steps}{length(Correct\ Action\ Script)}$$

$$F1\ Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

4. **実行率**：生成したアクションスクリプトのうち実行可能なアクションスクリプトの割合

- 評価方法

VH データセットの 1 行目（アクティビティ）を入力として、行動説明文のアクションスクリプトを出力する。VH データセットの 2 行目（行動説明文）と 4 行目以降（アクションスクリプト）を正解として評価を行う。

行動説明文の評価には ROUGE-1 を評価指標として用いる。また、アクションスクリプトの評価には LCSscore, Corectness, 実行率の 3 つを用いる。

アクティビティが同じ名前でも複数個（内容の違う）データが存在する場合、それぞれのデータに対して評価を行い、最も高かった値を評価値とする。

- 評価結果

表 2 に評価結果を示している。項目の「例の種類」は、提案手法の Comp.1 での学習データ（行動説明文）の与え方を示す。Desc. は VH データセットの 2 行目の行動の説明文、Script は VH データセットの 4 行目以降のアクションスクリプトを行動説明文に変換したものを与える。

「例の選択方法」は Few-shot での例の選択の仕方であり、ランダムに選択する方法と、入力となるアクティビティとの類似度の高いアクティビティのデータを選択する 2 つの方法がある。また、Fine-tuning では訓練用とテスト用で 5:5 に分割して評価を行っており、いずれも入力となるアクティビティと同じアクティビティを学習データとすることはしない。

※ より詳細な実験方法や結果については[\[青山+ 2023\]](#)を参照。

手法	例の種類	例の数	例の選択方法	ROUGE	LCSscore	Correctness			実行率 (%)
						Precision	Recall	F1 Score	
Zero-shot	-	-	-	0.33	0.32	0.38	0.61	0.44	96.5
Few-shot	Desc.	2	Random	0.45	0.34	0.48	0.54	0.46	100.0
		5	Random	0.48	0.34	0.45	0.54	0.46	100.0
		10	Random	0.51	0.40	0.52	0.57	0.51	100.0
	Script	2	Random	0.51	0.45	0.54	0.70	0.57	100.0
		5	Random	0.53	0.48	0.55	0.72	0.59	100.0
		10	Random	0.54	0.57	0.72	0.76	0.72	100.0
	Desc.	10	Sim.(Top)	0.52	0.42	0.62	0.60	0.57	100.0
	Script	10	Sim.(Top)	0.56	0.60	0.77	0.79	0.76	100.0
Fine-tuning	Decs.	-	-	0.44	0.46	0.70	0.70	0.66	100.0
	Script	-	-	0.37	0.45	0.63	0.72	0.64	100.0

表 2. 評価結果 ※ Sim.(Top)は類似度の高いアクティビティを選択する方法

#### - 構築したナレッジグラフの評価

上記の評価実験で評価値の良かった生成方法（手法：Few-shot, 例の種類：Script, 例の数：10, 例の選択方法：Sim.(Top)) で生成したデータの中から違った系統のアクティビティを6つ選択し, VirtualHome2KG を用いてオブジェクトの 3D 座標とアクションの実行時間を記録せずに状態変化のみをシミュレーションして生成したナレッジグラフと, 3D 座標と実行時間も記録して生成したナレッジグラフの2種類（計12個）を構築した.

構築したナレッジグラフは, 例えばナレッジグラフ推論チャレンジ【実社会版】2022 のチャレンジタスクである家庭内で起こり得る危険な状況の検出などの家庭内に関するタスクで利用できる.

#### - 自動でアクションスクリプトを構築する利点

アクションスクリプトを構築するためには, VH で実行可能なアクションとその制約, 環境の情報（部屋のレイアウト, オブジェクトの位置やタイプなど）をある程度理解する必要がある. また, データのバリエーションを増やすためには提供されるシーン1から7の環境の情報を熟知しなければならない. そのため, 手作業でアクションスクリプトを作成することは非効率である.

手作業での構築と比較して提案した手法は, アクティビティであれば簡単に記述することができるため, 効率的にアクションスクリプトを生成することができる. また, アクティビティを Chat-GPT などに大量生成させてしまえばアクションスクリプト生成の効率が向上すると考えられる.

## 6. ナレッジグラフの構築に利用したプログラム（オプション）

## 7. 資料の共有について

### 応募フォーム

- 公開の可否：
  - ☐ 公開してよい
  - ☐ 非公開とする

### 応募したナレッジグラフ

- 公開の可否：
    - ☐ 公開してよい
    - ☐ 非公開とする
  - 公開形式：
    - ☐ ナレッジグラフ推論チャレンジのサイトで公開
    - ☐ 独自のサイトで公開してリンクを希望
- 公開先 URL（※）：<https://github.com/JinAoyama/kgrc2023>

### 応募したプログラム等

- 公開の可否：
  - ☐ 公開してよい
  - ☐ 非公開とする
- 公開形式：
  - ☐ ナレッジグラフ推論チャレンジのサイトで公開
  - ☐ 独自のサイトで公開してリンクを希望