

LLMを活用した生活行動ナレッジグラフの構築

ナレッジグラフ推論チャレンジ2023
～生成AI時代のナレッジグラフ構築技術～

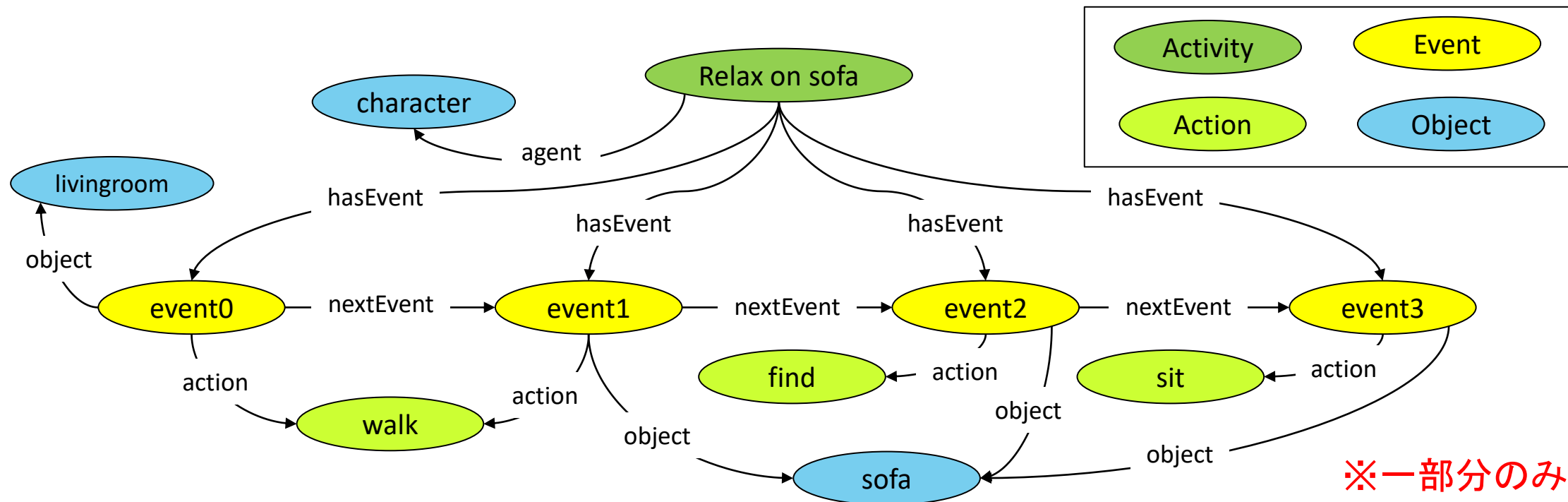
所属： 青山学院大学 森田研究室

氏名： 青山 仁

応募部門： 一般部門

構築したナレッジグラフ

- 日常生活における行動を表現したナレッジグラフ
 - 「[ナレッジグラフ推論チャレンジ【実社会版】2022](#)」で提供されているナレッジグラフ
 - 「イベント」とそれに対応する「行動」、「対象物」、「エージェント」、「場面間の関係」などで表現される



ナレッジグラフの構築手法 ①

●VirtualHome2KG

- 家庭用シミュレータ VirtualHome (VH) を用いて、「ソファで寛ぐ」、「テレビを見る」などの日常的な行動 (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

→ VHで日常的な行動をシミュレーションできれば、ナレッジグラフが構築可能

●VHでのシミュレーション実行方法

- 「アクション」、「オブジェクト」、「オブジェクトのID」の組のシーケンス表現 (アクションスクリプト) を用いる

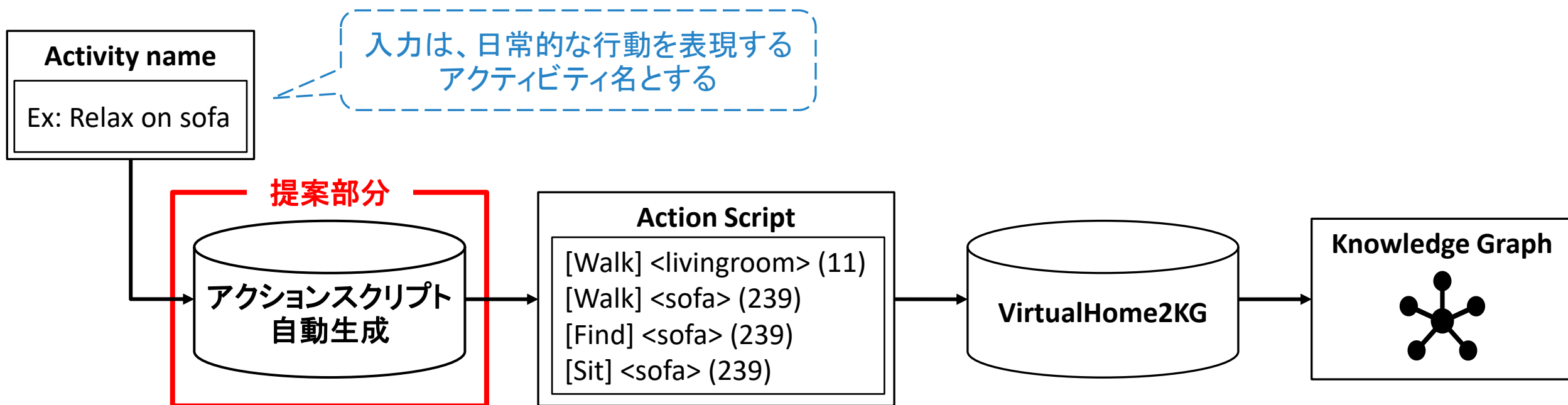
→ アクションスクリプトを自動生成できれば、ナレッジグラフの自動構築が可能

Action Script

```
[Walk] <livingroom> (11)
[Walk] <sofa> (239)
[Find] <sofa> (239)
[Sit] <sofa> (239)
```

ナレッジグラフの構築手法 ②

- LLM (GPT-3.5-Turbo) を利用したアクションスクリプト自動生成手法の提案をする
= LLM を利用したナレッジグラフ構築手法の提案



利用したデータ

●VHデータセット

- アクティビティ名, 行動説明文, アクションスクリプトをクラウドソーシングにより収集 [Puig+ 2018]

Relax on sofa ← アクティビティ名 (抽象的タスク記述)

I turn on the tv and sit in the sofa. I watch the tv. ← 行動説明文

[WALK] <livingroom> (1)

[WALK] <tv> (1)

[SWITCHON] <tv> (1)

[WALK] <sofa> (1)

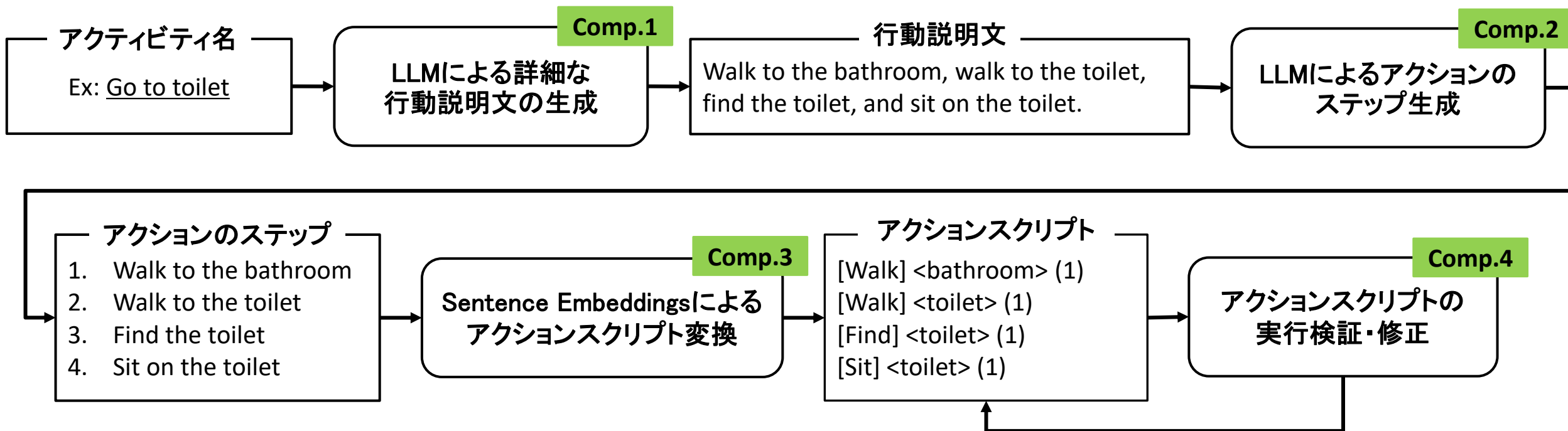
[SIT] <sofa> (1)

[TURNTO] <tv> (1)

[WATCH] <tv> (1)

← アクションスクリプト

アクションスクリプト自動生成の概要



Comp.1

LLMによる詳細な行動説明文の生成 ①

- LLMを用いて、アクティビティから詳細な行動説明文を生成する

プロンプト

user role

Describe the details of the following activity: {activity} #activityは変数

system role

1. You can use only the following actions in the description:
Walk, Find, Sit, ... #VH上で実行可能なアクション
2. You can use only the following objects and rooms in the description:
livingroom, tv, sofa, ... #環境の情報(環境内に存在する物または部屋)

Comp.1

LLMによる詳細な行動説明文の生成 ②

- 方法1: Zero-shotで生成
- 方法2: Few-shotで生成 → VHデータセットを活用

アクティビティに対応する
「行動説明文」or「アクションスクリプトを変換した説明文」
を例として与える2種類の方法を提案

Question

Describe the details of the following activity: **Relax on sofa**

Answer: Pattern1

I turn on the tv and sit in the sofa. #2行目の行動説明文

Answer: Pattern2

I walk to livingroom, walk to the tv, switch on the tv, ... #4行目以降のアクションスクリプトを説明文にする

VHデータセット

Relax on sofa ← **アクティビティ**

I turn on the tv and sit in the sofa. I watch the tv.

行動説明文

[WALK] <livingroom> (1)

[WALK] <tv> (1)

[SWITCHON] <tv> (1)

[WALK] <sofa> (1)

[SIT] <sofa> (1)

[TURNTO] <tv> (1)

[WATCH] <tv> (1)

アクションスクリプト

Comp.2

LLMによるアクションのステップ生成

- LLMを用いて、行動説明文からアクションのステップを生成する
→ 生成したアクションのステップはリスト形式で保存する

プロンプト

user role

Rewrite the following description as action steps: {description} #descriptionは変数

system role

1. You can use only the following actions in the description:
Walk, Find, Sit, ... #VH上で実行可能なアクション
2. You can use only the following objects and rooms in the description:
livingroom, tv, sofa, ... #環境の情報(環境内に存在する物または部屋)

Comp.3

Sentence Embeddingsによるアクションスクリプト変換

- Hash Tableの作成
 - VH上で実行可能な行動と、環境内に存在する物または部屋から成る得る行動文を Key として作成する
 - 行動文(Key)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
 - 生成したアクションのステップに対して、作成した辞書内で類似度の行動文(Key)を選択する
 - 選択した行動文(Key)に対応するアクションスクリプト(Value)に変換する

アクションスクリプト変換のための Hash Table

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

※ 類似度の計算には、Sentence Transformersの技術を用いる

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
 - VHシミュレータを用いて実行可能か検証する
 - 実行不可であった場合、原因となるアクションのステップを特定し、エラーメッセージを出力する
- アクションスクリプトの修正
 - エラーメッセージの内容から、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ある場合、対象オブジェクトに近づくアクション(Walk)を追加する
 - その他の原因の場合、失敗したアクションステップを削除する

評価概要

- 構築したナレッジグラフが、入力となるアクティビティを表現できているかを評価する
 - 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
 - 生成したアクションスクリプトの評価をする

- アクションスクリプト自動生成手法の評価
 - 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**(妥当かどうか)を正解データ(VHデータセット)を用いて評価
 - 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**(実行可能かどうか)

- 評価用データ: VHデータセット(197個)

評価指標

● LCSscore

- 行動の順序を考慮して一致度を計算する
- LCS (Longest Common Subsequence) をもとに以下のように定義

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

● 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

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

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

● 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

評価手順

1. VHデータセットのアクティビティ名を入力として、アクションスクリプトを生成する
2. 入力とするアクティビティに対応するアクションスクリプトを正解データとして、LCSscoreとCorrectnessを評価する

※ アクティビティに対応するデータが複数種類存在する場合、それぞれのデータに対して評価を行い、最高値を最終的な評価値とする

3. VHシミュレータを用いて、生成したアクションスクリプトの実行率を評価する

Few-shot学習での例の選択方法

1. ランダムに選択
2. 入力のアクティビティと類似度の高いアクティビティのデータを選択(入力と同じアクティビティのデータは選択しない)

以下、共通のルール

- 選択するデータの個数は10個
- アクティビティに対応するデータが複数個存在する場合、その中からランダムに1つ選択する
- アクティビティに重複がないように例を選択

評価結果

手法	例の種類	選択方法	LCSscore	Correctness			実行率 (%)
				Precision	Recall	F1 Score	
Zero-shot	-	-	0.32	0.38	0.61	0.44	96.5
Few-shot	Description	Random	0.40	0.52	0.57	0.51	100.0
		Sim.(Top)	0.42	0.62	0.60	0.57	100.0
	Script	Random	0.57	0.72	0.76	0.72	100.0
		Sim.(Top)	0.60	0.77	0.79	0.76	100.0

付録

生成したアクションスクリプト、構築したナレッジグラフの公開先(一部)

URL: <https://github.com/JinAoyama/kgrc2023>

QR:

