

Scene Graph Generation (シーングラフ生成)

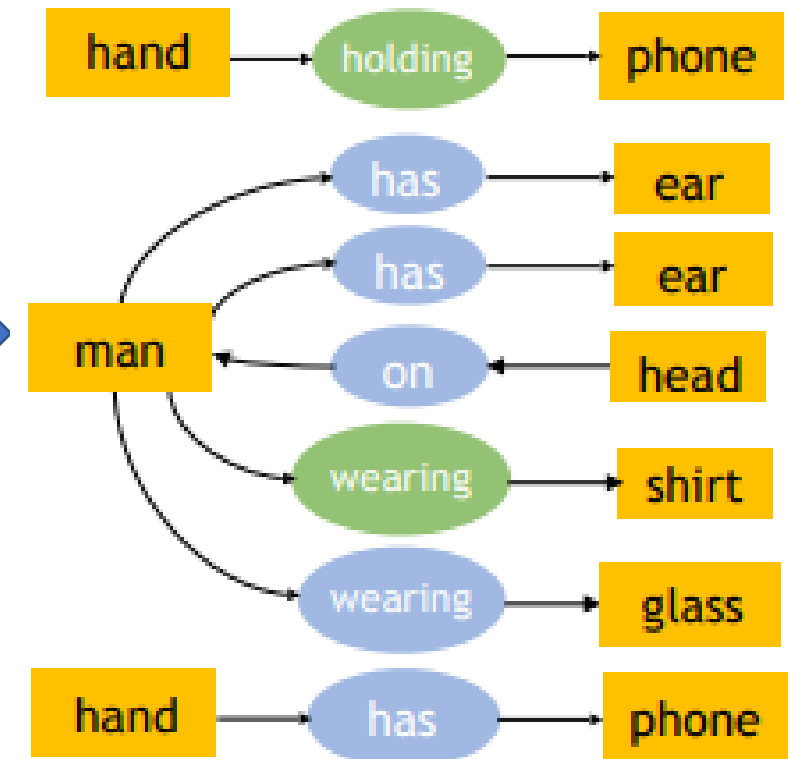
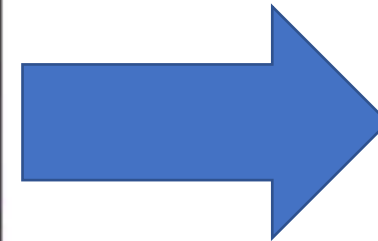
～動画からの物体と物体間の関係抽出～

202302010

はじめに

• Scene Graph Generation(SGG)とは?

- 画像を入力し物体検出、検出したオブジェクト間の位置関係と意味情報をグラフに表現するタスク



SGGの応用例

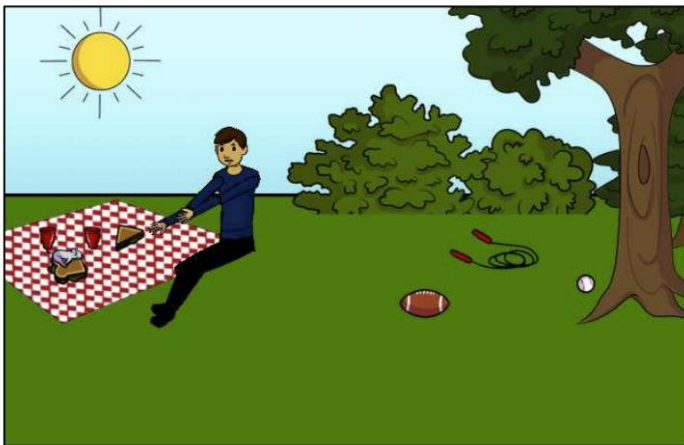
- SGGは主にVisual Question Answering(VQA)に応用されている



What color are her eyes?
What is the mustache made of?



How many slices of pizza are there?
Is this a vegetarian pizza?



Is this person expecting company?
What is just under the tree?



Does it appear to be rainy?
Does this person have 20/20 vision?

VQAの例(左図):画像と質問が与えられた時、
自然言語の答えを提供するタスク



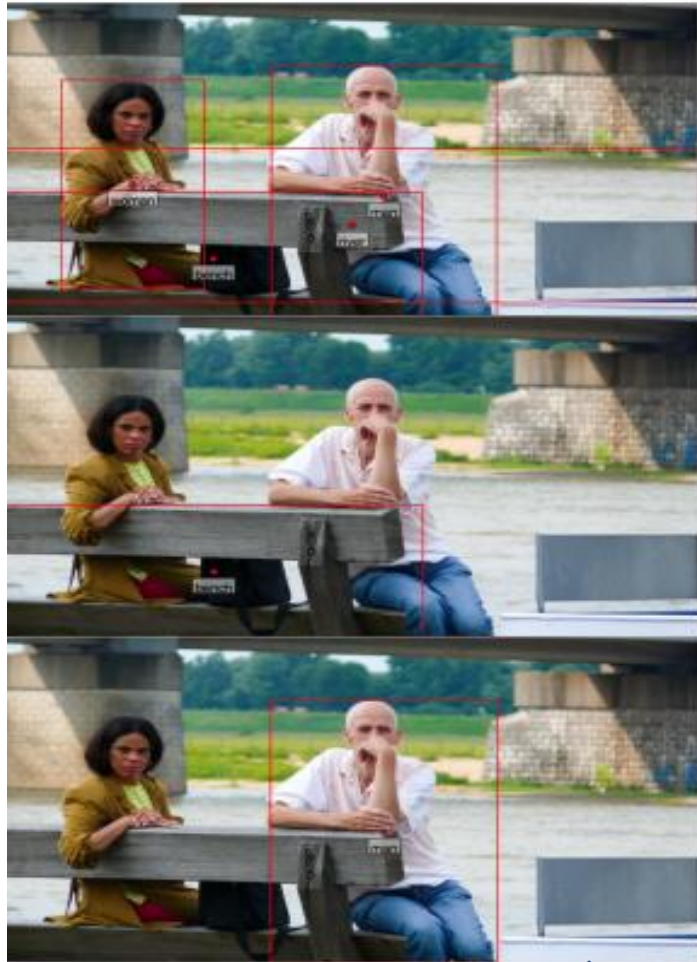
VQAのタスクをこなすには、
**画像内にはどのようなオブジェクトがあり、
オブジェクト間にどのような関係があるのか、
画像内から読み取る視覚的な理解が重要**



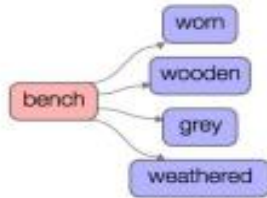
Scene Graph Generation(SGG)のように
オブジェクトの位置、関係を検出する
タスクがVQAタスクの重要な基礎となる

現在のSGGで提供されているデータセット

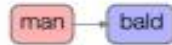
- 現在SGGに特化したDataSetが作成されてことにより、盛んに研究されている。



A man and a woman sit on a park bench along a river.

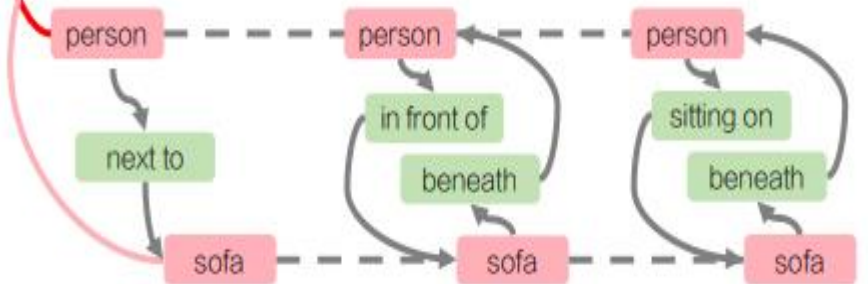


Park bench is made of gray weathered wood



The man is almost bald

ActionGenome: SceneGraph表現に時間の要素を追加



Spatio-temporal scene graphs

VisualGenome: オブジェクト間の関係をSceneGraphで表現

Scene Graph Generation学習データ

画像



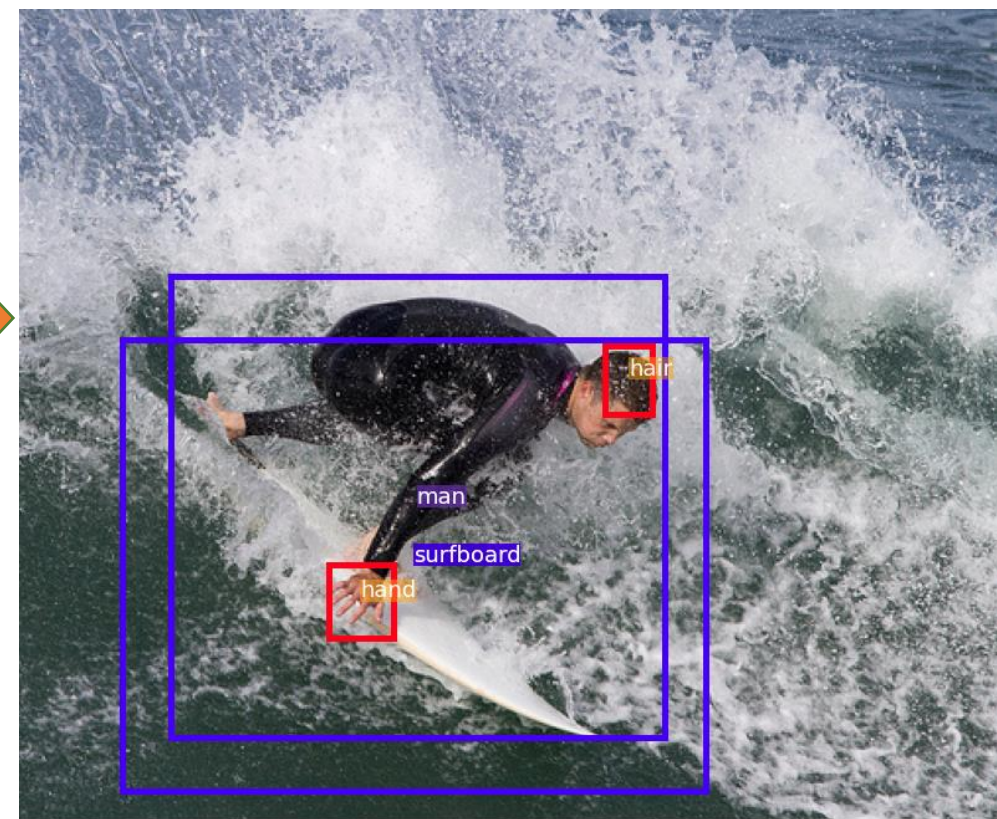
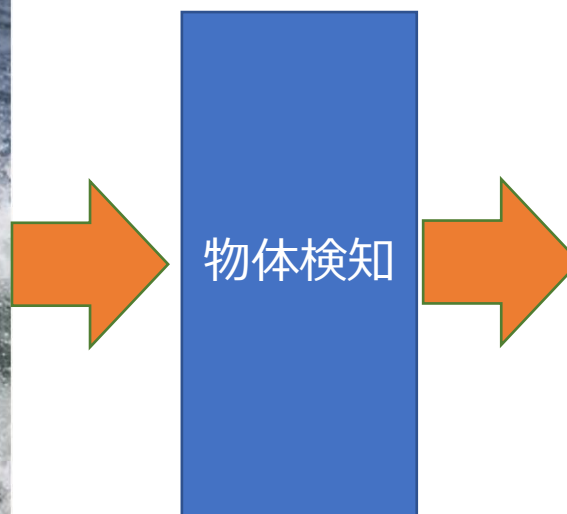
シーングラフ

```
{'1.jpg': [{'predicate': 48,
  'subject': {'category': 77, 'bbox': [246, 507, 236, 295]},
  'object': {'category': 119, 'bbox': [487, 513, 240, 292]}},
{'predicate': 19,
  'subject': {'category': 77, 'bbox': [246, 507, 236, 295]},
  'object': {'category': 110, 'bbox': [285, 388, 239, 293]}},
'2.jpg': [{'predicate': 19,
  'subject': {'category': 21, 'bbox': [0, 414, 567, 795]},
  'object': {'category': 144, 'bbox': [66, 231, 646, 740]}},
{'predicate': 19,
  'subject': {'category': 21, 'bbox': [0, 414, 567, 795]},
  'object': {'category': 144, 'bbox': [68, 229, 744, 797]}},
. . . . .
}]
```

※'predicate': xx、'category': xxは、
別ファイルpredicates.json、objects.jsonに対応する番号

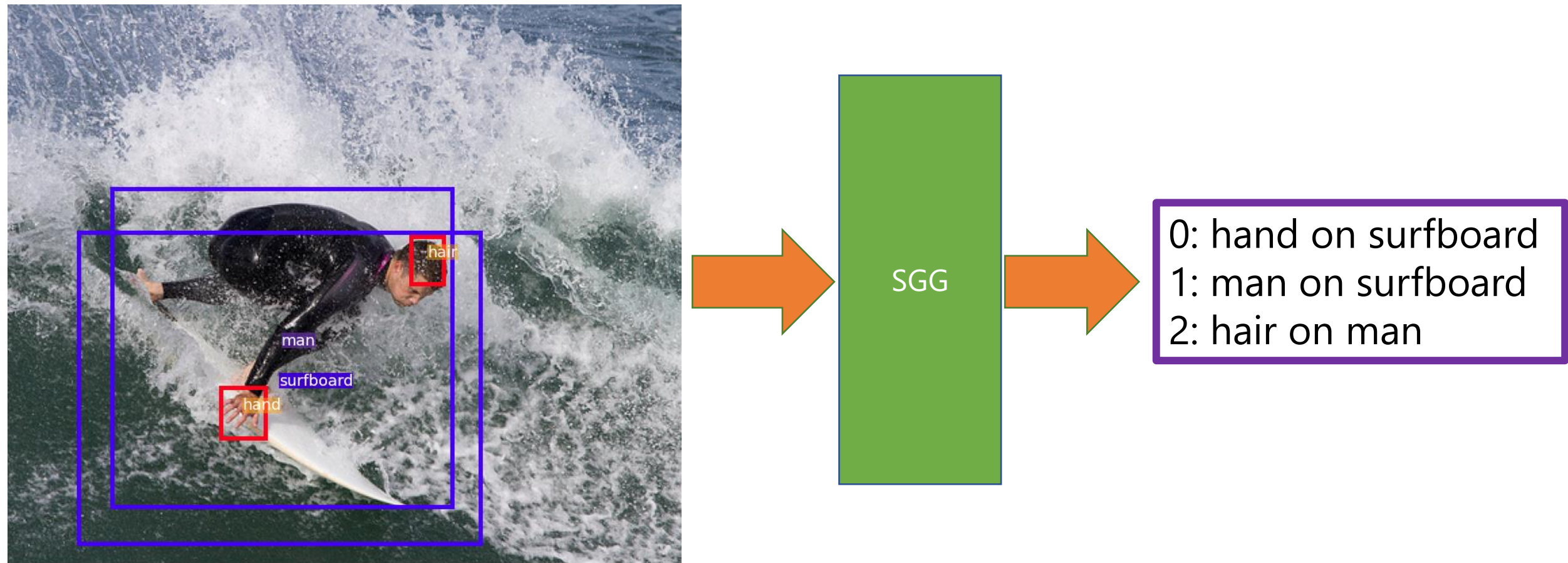
Scene Graph Generation

画像を入力し、オブジェクトを検出する



Scene Graph Generation

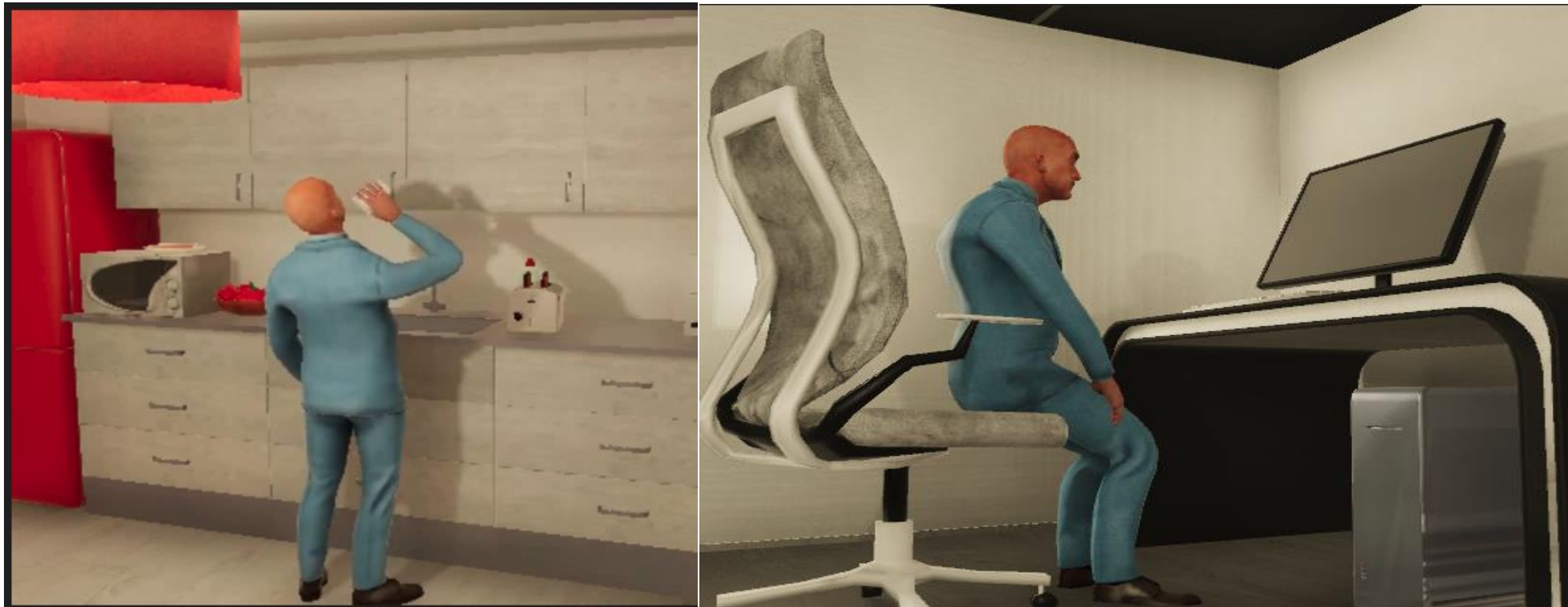
検出したオブジェクト間の、関係を検出する



SGGをVirtual Homeに応用

●SGGをVirtual Homeに応用

- 仮想空間上でアバターによる屋内行動の生成が可能なVirtualHomeをSGGに応用
- アバターに仮想空間で日常的な行動をとらせて、
周りのオブジェクトとの相互関係の情報を入手できる



上図:VirtualHomeのDrinkingアクション(左)、Sitting on(右)

SGGをVirtual Homeに応用

