



VirtualHome

Eneko Alaminos and Aitor Domec



VirtualHome

Simulating Household Activities via Programs



Make a
margarita

Store groceries

Clean

...

Make the bed

Make a margarita

Prepare dinner

Wash dishes

Set up the table

Play Bingo with me

Take me pee



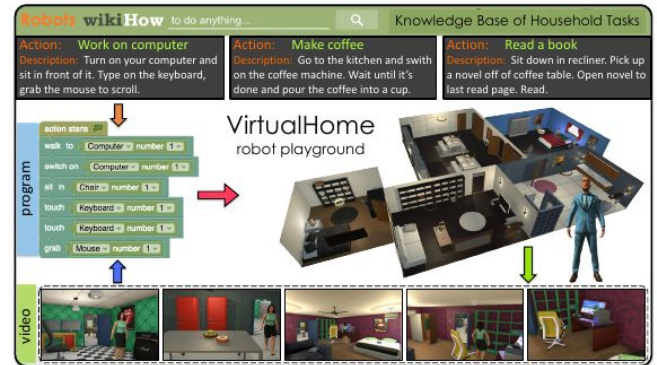
Goals

- ❑ Creation of a knowledge base to train agents.
- ❑ Transcription from language and video to program.

```
action starts
walk to living_room number 1
walk to book number 1
grab book number 1
walk to sofa number 1
lie in sofa number 1
read book number 1
```

```
action starts
walk to living_room number 1
walk to remote_control number 1
find remote_control number 1
grab remote_control number 1
find television number 1
switch on television number 1
turn to television number 1
watch television number 1
```

```
action starts
walk to kitchen number 1
walk to coffee_pot number 1
find coffee_pot number 1
grab coffee_pot number 1
find coffee number 1
grab coffee number 1
open coffee_maker number 1
put coffee number 1 in coffee_maker number 1
drop coffee_pot number 1
find coffee_maker number 1
switch on coffee_maker number 1
switch off coffee_maker number 1
drink from coffee number 1
```





Parts of VirtualHome

1. Dataset
 - a. Obtention
 - b. Description
2. Simulator
 - a. Made in Unity3D
 - b. Python API to communicate with it.
3. A model to “translate” from description or videos to programs.

Dataset Obtention


- Data collection via crowdsourcing
 - -AMT & Upwork-
- 1. Data description
 - a. Household activities
 - b. 8 possible scenes
- 2. Translate into programs
- 3. Double-check the data




Dataset Description

- Program = Sequence of instructions
 - 77 possible blocks
 - 75 atomic actions
 - 308 objects
- 2821 programs, 576 tasks


Read a book

```
action starts   
walk to living_room number 1  
walk to book number 1  
grab book number 1  
walk to sofa number 1  
lie in sofa number 1  
read book number 1
```

Watch TV

```
action starts   
walk to living_room number 1  
walk to remote_control number 1  
find remote_control number 1  
grab remote_control number 1  
find television number 1  
switch on television number 1  
turn to television number 1  
watch television number 1
```

Make Coffee

```
action starts   
walk to kitchen number 1  
walk to coffee_pot number 1  
find coffee_pot number 1  
grab coffee_pot number 1  
find coffee number 1  
grab coffee number 1  
find coffe_maker number 1  
open coffe_maker number 1  
put coffee number 1 in/on coffe_maker number 1  
drop coffee_pot number 1  
find coffe_maker number 1  
switch on coffe_maker number 1  
switch off coffe_maker number 1  
drink from coffee number 1
```

Simulator



Example 1 (NavMesh)

[Find] <Shower> (1)
[Walk] <Shower> (1)



Example 2 (Inverse Kinematics)

[Walk] <chair> (1)

[Find] <cat> (1)

[Grab] <cat> (1)

[Walk] <microwave> (1)

[Open] <microwave> (1)



Example 3

[Find] <plate> (1)
[Grab] <plate> (1)
[Find] <microwave> (1)
[Open] <microwave> (1)
[PutIn] <plate> (1) <microwave> (1)
[SwitchOn] <microwave> (1)
[Find] <milk> (1)
[Grab] <milk> (1)
[Drink] <milk> (1)
[Find] <chair> (1)
[Sit] <chair> (1)




From Description or Videos to Programs





Conclusion

- ❑ **Goal:** Description or Videos  Programs
- ❑ 3D Environment in Unity3D.
- ❑ Presented a model to achieve the goal.

Python API

- ❑ They provide a Python API to communicate with the simulator.
- ❑ It is not perfect, but it easy to use.
- ❑ Problems at generating videos.