Interactive Question Answering

Language learning using text-based games

Edan Toledo | Gregory Furman | Roy Cohen | Dr. Jan Buys | Dr. Jonathan Shock

Project Overview

We investigate three methods for training question-answering agents in text-based environments. Through interaction and observation, these agents were required to answer questions about objects located within dynamically created text games in an interactive question-answering (IQA) setting.

Aims

- Improve upon previous question-answering benchmarks.
- Highlight the inferior sample efficiency of prior methods.

Environment

- **Games** are procedurally generated with each consisting of a number of rooms containing objects.
- **Actions** are natural language commands that consist of an action, modifier, and object i.e “eat red toast”
- **Map** types indicate the number of rooms in an environment. **Fixed** maps have 6 rooms. **Random** maps have between 2 and 12 rooms.

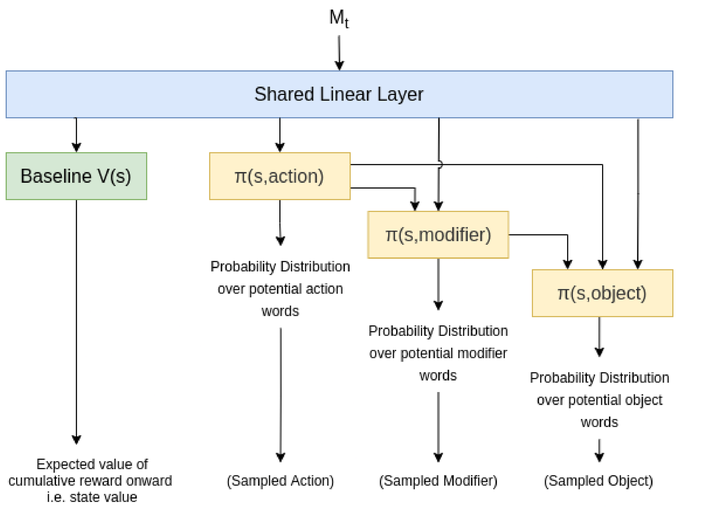
Game	You find yourself in a backyard. You make out a patio table. You see a patio chair. It’s stylish. In a corner you can see a BBQ. There’s an open screen door leading south.
Quest	Where is the diced potato?
Agent	>**Qapla’**
Game	That’s not a verb I recognize.
Agent	>**eat ghargh**
Game	You can’t see any such thing.
Agent	>**go south**
Game	You’ve just entered a kitchen. You can see an open fridge here. It is empty! You see a closed oven. Look over there, it’s a counter. On the counter you can see a diced potato, a red ghargh and a knife. There’s an open door is leading north.
Agent	>**wait**
Agent	>**answer: kitchen**

Approaches

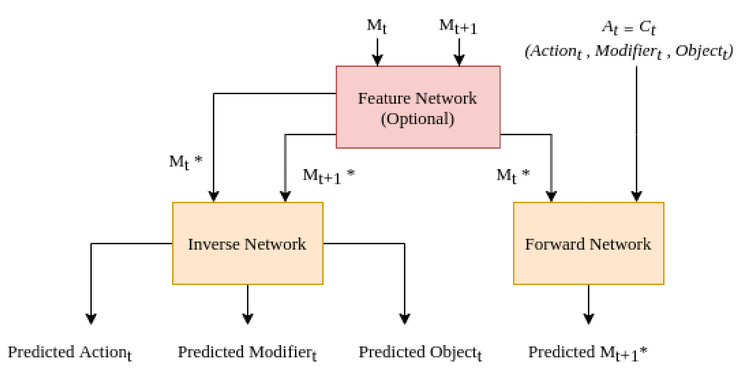
Policy-Based

We investigated the effectiveness of a policy-based reinforcement learning method along with a learned environment dynamics model to promote semantic encoding.

The learned environment dynamics model aims to give the agent an understanding of the enviornment dynamics.



Policy and Value network architecture



Environment Dynamics architecture

- Higher training performance.
- Better Generalisation performance.
- Greater Sample efficiency.
- Environment dynamics model increases generalisation performance when access to more than 1 unique game is not available.

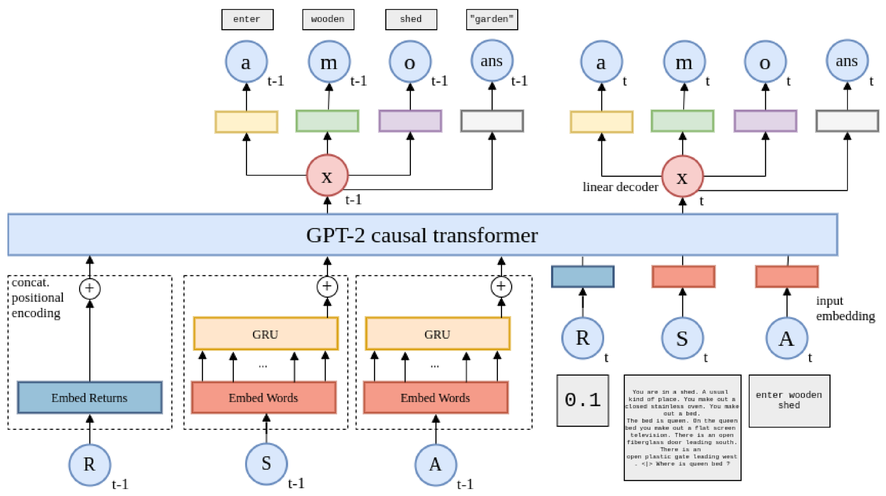
1 Fixed Game			
Model	Location	Existence	Attribute
Best Performing Baseline	0.164	0.628	**0.516**
Policy-based	0.168	0.584	0.514
Policy-based w/ Environment Dynamics	**0.182**	**0.630**	0.494
500 Fixed Games			
Model	Location	Existence	Attribute
Best Performing Baseline	0.224	0.674	**0.534**
Policy-based	**0.948**	**0.948**	0.466

Question-answering accuracy achieved by the policy-based agent.*

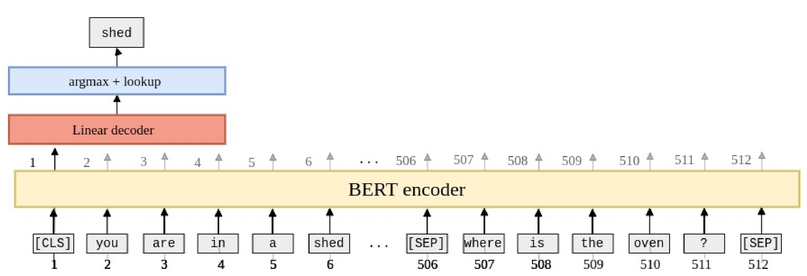
Sequence-Modelling

By framing IQA as an offline sequence modelling problem, we investigate the applicability of the novel Decision Transformer (DT) architecture. A sequence consists of states, actions, and rewards corresponding to each episodic timestep t.

A causally masked GPT-2 Transformer for action generation along with an answer prediction head for QA was used. Additionally, we trained a BERT model for question-answering to be used in tandem with the DT’s command generation heads (DT-BERT).



Decision Transformer architecture



BERT QA model architecture

- Greater sample efficiency.
- Increased navigational and QA capabilites than several RL baselines.
- Decreased training time.

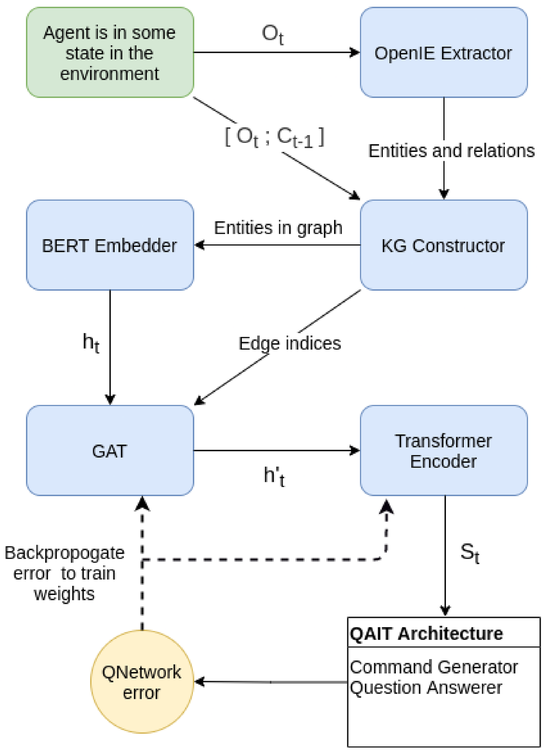
500 Random Games			
Model	Location	Existence	Attribute
Best Performing Baseline	0.222	0.678	0.530
DT	0.104	**0.722**	0.526
DT-BERT	**0.270**	0.654	**0.538**

*Question-answering accuracy achieved by the DT & DT-BERT.**

Graph Attention Network

This method explores the use of Graph Attention Networks (GATs) in providing an RL agent with some contextual understanding of the environment with which it inhabits.

Context is provided by embedding specific details about an environment into a knowledge graph, which the GAT utilises to provide additional context to an agent.



IQA-GAT architecture

- GATs can increase the question answering accuracy of an agent when training on data containing high variability.
- Improved agent generalisation abilities.

Unlimited Random Games			
Model	Location	Existence	Attribute
DQN	0.188	0.668	0.506
DQN+GAT	**0.212**	**0.708**	**0.512**

Question-answering accuracy achieved by the DQN using the GAT.*

Results & Conclusions

**tables only show subset of full results, which are available in the paper.*

University of Cape Town
Computer Science Department
www.cs.uct.ac.za
Email: dept@cs.uct.ac.za
Telephone: +27 21 650 2663

UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA • UNIVERSITEIT VAN KAAPSTAD

Authors
Edan Toledo (Policy-Based) < *tldeda001@myuct.ac.za* >
Gregory Furman (Sequence-Modelling) < *frmgre001@myuct.ac.za* >
Roy Cohen (Graph Attention Network) < *chnroy002@myuct.ac.za* >

Supervisors
Dr. Jan Buys < *jbuys@cs.uct.ac.za* >
Dr. Jonathan Shock (co) < *jonathan.shock@uct.ac.za* >