

A Appendix

We would first like to note the presence of a complementary dataset of observation-action pairs created by humans on the ClubFloyd online Interactive Narrative forum.⁶ This dataset appears in both Ammanabrolu and Hausknecht [2020] and Yao et al. [2020] with the latter using it to tune a GPT-2 model for valid action prediction. To prevent data leakage from human transcripts to our test games, we do *not* use this dataset to pre-train or tune our models.

The rest of this Appendix first provides additional samples for the dataset for qualitative purposes and then provides training details for the baseline models.

A.1 Dataset

The games used in the Jericho suite and here are all open sourced freeware. The walkthroughs required to create the oracle agents for the collection of data for the games were drawn from various sources on the internet and errors were corrected manually. We provide our data at <https://github.com/JerichoWorld/JerichoWorld> under an MIT license. We provide 3 samples drawn from different games in the full dataset to help the readers better understand the diversity of text there.

Game: 905
Location:
 Bedroom (in bed)
 This bedroom is extremely spare, with dirty laundry scattered haphazardly all over the floor. Cleaner clothing can be found in the dresser. A bathroom lies to the south, while a door to the east leads to the living room.

 On the end table are a telephone, a wallet and some keys.

 The phone rings.
Observation: You take off the gold watch. The phone rings.
Inventory:
 You are carrying:
 some soiled clothing (being worn)
 a gold watch
Prev Act: take off watch
Inventory Objects:
 gold watch: Apparently it's 9:07. The phone rings.
 soiled clothing: These clothes are a lost cause, sad to say no amount of laundering is going to get these stains out.
Inventory Attributes:
 watch: animate, equip
 clothing: animate, equip
Surrounding Objects:
 phone: An ordinary telephone, notable chiefly for being fifteen or twenty years old.
 keys: House keys, car keys, they're all on the same chain.
 end table: A small end table, oak veneer over plywood.
 living room: The living room lies to the east.
 dirty dresser: Just a simple dresser.
 laundry: Shirts, pants, the usual.
 floor, east, south: You see nothing unexpected in that direction. The phone rings.
 wallet: It's a brown leather wallet.
 door: Just a regular door.
Surrounding Attributes:
 keys: animate, equip
 wallet: animate, equip
Graph: [you, have, gold watch], [you, in, bed], [you, have, soiled clothing]
Valid Actions: take phone, get up, take off clothing, take off watch, take keys, close door, take wallet, close door, put clothing down, put watch down, put clothing on table, open wallet, put watch down, put clothing on phone, put watch on table, put gold on phone, look under bed

Game: deephome
Location:
 Secret Entrance
 This is a rather dark and small room, having only two exits, back north the way you came, from the ancestral homes of Tana, or through the heavily barred wooden door before you that leads southwest and inward to the abandoned Deephome, abode of the Dwarves in Telleen. It has been three hundred years since your people lived here.

 The heavy door stands open, admitting you into Deephome.
Observation: As you touch the finely etched symbol, you hear a click and a whirl. Then the door swings open before you, opening into the abandoned city of Deephome. Your score has just gone up by five points.
Inventory:

⁶http://www.allthingsjacq.com/interactive_fiction.html

You are carrying:
 King's Order
 a lantern (providing light)
 Prev Act: push mountain
 Inventory Objects:
 lantern: This is an old and trusty (not rusty) lantern that has been in your family for centuries. It has yet to shut off at an inopportune moment. However, there is a saying in your family..."That lantern is bound to go off at an inopportune time sometime!"
 order: The note reads: "Reclaimer: You have the esteemed duty to return to our Mountain Kingdom of Deephome and prepare it for our return. There are several things a Reclaimer must do: 1. Restore Power to the City 2. Restore Water to the city. 3. Visit each location and make sure it is safe, a quick appraisal should be sufficient. 4. Open the City Gates once more. 5. MOST IMPORTANT: Make sure the city is SAFE to return to. May the Peace of Kraxis go with you King Derash of the Mountain Tana, the year 782 SK."
 Inventory Attributes:
 lantern: equip
 Surrounding Objects:
 southwest: You see nothing special about the southwest wall.
 house: It is the typical human house, maybe two stories. It is etched into the wood.
 wooden door: This door is made of thick and sturdy wood. It has three symbols on it, a tree, a house, and a mountain.
 symbols: On the door there are pictures of a mountain, a tree, and a house.
 tree: The tree symbol looks as if it were etched into the wood.
 mountain: The mountain looks mighty, a high peak among the clouds. It is etched into the wood.
 Surrounding Attributes:
 door: unlockable
 symbols: unlock
 Graph: [symbols, in, Secret Entrance], [wooden door, in, Secret Entrance], [ground, in, Secret Entrance], [you, in, Secret Entrance], [house, in, Secret Entrance], [Kraxis, in, Secret Entrance], [you, have, lantern], [mountain, in, Secret Entrance], [you, have, "Kings Order"], [tree, in, Secret Entrance]
 Valid Actions: say manaz, push mountain, close wooden, get in southwest, put light down, put order down

 Game: reverb
 Location:
 Behind the Counter
 You are behind the counter at "Mr. Tasty's Pizza Parlor". To the southwest is the rest of the restaurant.

 On the counter is a large pizza box (which is closed).

 You can see a handwritten note here.
 Observation: You put the large pizza box on the counter.
 Inventory: You are carrying nothing.
 Prev Act: push large to counter
 Inventory Objects:
 Inventory Attributes:
 Surrounding Objects:
 southwest: You see nothing special about the southwest wall.
 handwritten note: The note reads: "Stanley, Don't forget to make your delivery to Mr. Calzone, located at the San Doppleton Courthouse. You're already on thin ice, kid. One more screwup and you can expect to be looking for a new job." The note is signed with the initials "RT". The paper is official "Mr. Tasty's" stationery with the name Bob "Tasty" Tasker and lots of balloons and smiley faces all over the border. Isn't that cute?
 large pizza box: It's a large, flat, greasy cardboard box. Hastily scrawled on the outside is the word "Calzone". Which is weird, because it's clearly a pizza.
 counter: It's a majorly boring counter which you're unfortunately very familiar with.
 Surrounding Attributes:
 handwritten note: indoor, readable
 large pizza box: indoor
 counter: indoor
 Graph: [metal file, in, large pizza], [you, in, Behind the Counter], [handwritten note, in, Behind the Counter], [large pizza, in, large pizza box], [counter, in, Behind the Counter], [large pizza box, in, counter]
 Valid Actions: get up, take note, take large, examine note, undo large, push note to southwest, push large to southwest, push note to counter, push large to counter

A.2 Baselines

The baseline models that are adapted from other works, i.e. the Rules and QA systems, are trained using hyperparameters and methodologies described in their respective works.

A.2.1 Rules

Following Ammanabrolu and Hausknecht [2020], the exact details regarding knowledge graph updates are found as follows. At every step, given the current state and possible attributes as context. The rest of the triples are extracted using OpenIE [Angeli et al., 2015].

- Linking the current room type (e.g. “Kitchen”, “Cellar”) to the items found in the room with the relation “has”, e.g. $\langle kitchen, has, lamp \rangle$
- All attribute information for each object is linked to the object with the relation “is”. e.g. $\langle egg, is, treasure \rangle$
- Linking all inventory objects with relation “have” to the “you” node, e.g. $\langle you, have, sword \rangle$
- Linking rooms with directions based on the action taken to move between the rooms, e.g. $\langle Behind\ House, east\ of, Forest \rangle$ after the action “go east” is taken to go from behind the house to the forest

A.2.2 Question-Answering

The QA models are trained on the SQuAD 2.0 [Rajpurkar et al., 2018], the Jericho-QA text game question answering dataset on the same set of training games as found in JerichoWorld, and then on JerichoWorld itself by formatting our dataset in the style of questions and answers when possible. Our dataset is formatted in the style of Jericho-QA by templating questions that ask about location, objects (including characters), and attributes. An example of a JerichoWorld dataset example converted to Jericho-QA format is seen below—though we would like to note that this removes much of the information present naturally within our dataset. All other model architecture and hyperparameter details are as seen in Ammanabrolu et al. [2020b].

```
Game: reverb
Location:
  Behind the Counter
  You are behind the counter at "Mr. Tasty's Pizza Parlor". To the southwest is the rest of the
  restaurant.

  On the counter is a large pizza box (which is closed).

  You can see a handwritten note here.
Observation: You put the large pizza box on the counter.
Inventory: You are carrying nothing.
Question: Where am I located? Answer: Behind the Counter
Question: What is here? Answer: large pizza box, handwritten note, southwest
Question: What do I have? Answer: nothing
Question: What attributes does handwritten note have? Answer: indoor, readable
Question: What attributes does southwest have? Answer: indoor
Question: What attributes does large pizza box have? Answer: indoor
```

A.2.3 Seq2Seq

For both tasks, models were trained until validation accuracy (picked to be a random 10% subset of the training data) did not improve for 5 epochs or 72 wall clock hours on a machine with 4 Nvidia GeForce RTX 2080 GPUs, three times with three random seeds. All models decode using beam search with a beam width of 15 at test time until the end-of-sequence tag is reached. The size of the decoding vocabulary for the action prediction task is 11056 and for the graph prediction task is 6985. Hyperparameters were not tuned and were taken from BART [Lewis et al., 2020].

Hyperparameter type	Value
Dictionary Tokenizer	Byte-pair encoding
Num. Encoder layers	6
Num. Decoder layers	6
Num. encoder and decoder attention heads	8
Feedforward network hidden size	4096
Input length	1024
Embedding size	768
Batch size	16
Dropout ratio	0.1
Gradient clip	1.0
Optimizer	Adam
Learning rate	1.0×10^{-3}

Table 4: Hyperparameters used to train the Seq2Seq model. It has a total of 232 million trainable parameters.

B Datasheet

We provide comprehensive documentation of the dataset based on Datasheets for Datasets [Gebru et al., 2018].

B.1 Motivation

For what purpose was the dataset created? Was there a specific task in mind? Was there a specific gap that needed to be filled? Please provide a description. We seek to create agents that exhibit human-like capabilities such as commonsense reasoning and natural language understanding in interactive and situated settings. In pursuit of this goal, we provide a dataset that enables the creation of learning agents that can build knowledge graph-based world models of interactive narratives.

Who created this dataset (e.g., which team, research group) and on behalf of which entity (e.g., company, institution, organization)? It was created by Prithviraj Ammanabrolu and Mark Riedl at the Georgia Institute of Technology.

Who funded the creation of the dataset? If there is an associated grant, please provide the name of the grantor and the grant name and number. It was funded by the US’s Defense Advanced Research Projects Agency (DARPA) as part of a fundamental science research grant Science of Artificial Intelligence and Learning for Open-world Novelty (SAIL-ON <https://www.darpa.mil/program/science-of-artificial-intelligence-and-learning-for-open-world-novelty>).

B.2 Composition

What do the instances that comprise the dataset represent (e.g., documents, photos, people, countries)? Are there multiple types of instances (e.g., movies, users, and ratings; people and interactions between them; nodes and edges)? Please provide a description. Each instance of our dataset takes the tuples of $\langle s_t, a_t, s_{t+1}, r_{t+1} \rangle$ where s_t and s_{t+1} are two subsequent states of a text game with a_t being the action used to transition states and r_{t+1} is the observed reward for some step t . Everything is in text. These are all collected from various text games and examples of instances are found in Appendix A.1.

How many instances are there in total (of each type, if appropriate)? The training data has 24198 mappings and is collected across 27 games in multiple genres and contains a further 7836 heldout instances over 9 additional games in the test set.

Does the dataset contain all possible instances or is it a sample (not necessarily random) of instances from a larger set? If the dataset is a sample, then what is the larger set? Is the sample representative of the larger set (e.g., geographic coverage)? If so, please describe how this representativeness was validated/verified. If it is not representative of the larger set, please describe why not (e.g., to cover a more diverse range of instances, because instances were withheld or unavailable). The dataset is a sample of the larger set of all possible states in each game. The samples are made to be biased towards states near the walkthroughs required to finish a game.

What data does each instance consist of? “Raw” data (e.g., unprocessed text or images) or features? In either case, please provide a description. Data is all in the form of text, either raw or in structured knowledge graph form.

Is there a label or target associated with each instance? If so, please provide a description. The data has multiple fields, depending on the tasks defined any of them can be used as labels. E.g. the knowledge graph prediction task has the graph field as the target.

Is any information missing from individual instances? If so, please provide a description, explaining why this information is missing (e.g., because it was unavailable). This does not include intentionally removed information, but might include, e.g., redacted text Not all games have human readable attributes for objects—when they do not, these are omitted by leaving the attributes fields blank. All other data is present for all instances.

Are relationships between individual instances made explicit (e.g., users’ movie ratings, social network links)? If so, please describe how these relationships are made explicit. Instances are grouped together by game through the game field.

Are there recommended data splits (e.g., training, development/validation, testing)? If so, please provide a description of these splits, explaining the rationale behind them. We provide a training split of 27 games, and a testing split of 9 games. These are selected on the basis of existing works and each split contains a diverse set of games in terms of genre.

Are there any errors, sources of noise, or redundancies in the dataset? If so, please provide a description.

Is the dataset self-contained, or does it link to or otherwise rely on external resources (e.g., websites, tweets, other datasets)? If it links to or relies on external resources, a) are there guarantees that they will exist, and remain constant, over time; b) are there official archival versions of the complete dataset (i.e., including the external resources as they existed at the time the dataset was created); c) are there any restrictions (e.g., licenses, fees) associated with any of the external resources that might apply to a future user? Please provide descriptions of all external resources and any restrictions associated with them, as well as links or other access points, as appropriate. The creation of the dataset depends on the Jericho framework <https://github.com/microsoft/jericho> but the archival versions themselves do not have any dependencies.

Does the dataset contain data that might be considered confidentiality, data that includes the content of individuals non-public communications)? If so, please provide a description. No, all data is part of games that are already public.

Does the dataset contain data that, if viewed directly, might be offensive, insulting, threatening, or might otherwise cause anxiety? If so, please describe why. The data is collected from games containing situations of non-normative language usage—describing situations that fictional characters may engage in that are potentially inappropriate, and on occasion impossible, for the real world such as running a troll through with a sword. Instances of such scenarios are mitigated by careful curation of the games that the data is collected from. The original Jericho framework [Hausknecht et al., 2020]—further verified by us in this work—uses a curated set of games found not to contain extreme examples of non-normative language usage. This is based on manual vetting and (existing) crowd-sourced reviews on the popular interactive narrative forum IFDB <https://ifdb.org/>.

B.3 Collection

How was the data associated with each instance acquired? Was the data directly observable (e.g., raw text, movie ratings), reported by subjects (e.g., survey responses), or indirectly inferred/derived from other data (e.g., part-of-speech tags, model-based guesses for age or language)? If data was reported by subjects or indirectly inferred/derived from other data, was the data validated/verified? If so, please describe how We build off the popular text game simulator Jericho [Hausknecht et al., 2020], we have constructed a dataset dubbed JerichoWorld that maps text game state observations to both the underlying ground truth knowledge graph representations of the game and the set of contextually relevant actions that can be performed in that state.

What mechanisms or procedures were used to collect the data (e.g., hardware apparatus or sensor, manual human curation, software program, software API)? How were these mechanisms or procedures validated? To collect the $\langle s_t, a_t, s_{t+1}, r_{t+1} \rangle$ tuples we implement a basic agent that explores the game along a trajectory corresponding to a *game walkthrough*. Game walkthroughs are texts describing the solutions to games, generally retrieved from the internet, but already part of the Jericho framework. Walkthroughs, however, only present one possible solution to a game and solve all the core puzzles required to complete a game with the maximum possible score. To achieve greater coverage of the game’s state space, our data collection agent stops off to explore by executing random valid actions for n steps before resetting to the walkthrough.

If the dataset is a sample from a larger set, what was the sampling strategy (e.g., deterministic, probabilistic with specific sampling probabilities)? Randomly sampled actions are based on a random seed in Python’s random package <https://docs.python.org/3/library/random.html>. We provide a seed and the specific package version.

Who was involved in the data collection process (e.g., students, crowdworkers, contractors) and how were they compensated (e.g., how much were crowdworkers paid)? Only the authors were involved, building on the contributions of the Jericho developers.

Over what timeframe was the data collected? Does this timeframe match the creation timeframe of the data associated with the instances (e.g., recent crawl of old news articles)? If not, please describe the timeframe in which the data associated with the instances was created. This dataset was developed over a period of 6 months, though the games used within date back to the 1970s.

Were any ethical review processes conducted (e.g., by an institutional review board)? If so, please provide a description of these review processes, including the outcomes, as well as a link or other access point to any supporting documentation. No human subjects were involved, no IRB process was undertaken.

B.4 Preprocessing

Was any preprocessing/cleaning/labeling of the data done (e.g., discretization or bucketing, tokenization, part-of-speech tagging, SIFT feature extraction, removal of instances, processing of missing values)? If so, please provide a description. If not, you may skip the remainder of the questions in this section. Games were decompiled to extract attributes and ground truth knowledge graphs, the creation script is provided in the GitHub repo.

Was the “raw” data saved in addition to the preprocessed/cleaned/labeled data (e.g., to support unanticipated future uses)? If so, please provide a link or other access point to the “raw” data. No, raw binary game states were not saved and were converted to human readable text.

Is the software used to preprocess/clean/label the instances available? If so, please provide a link or other access point. Games were decompiled to extract attributes and ground truth knowledge graphs, the creation script will be provided in the GitHub repository.

B.5 Uses

Has the dataset been used for any tasks already? If so, please provide a description. No.

Is there a repository that links to any or all papers or systems that use the dataset? If so, please provide a link or other access point. No.

What (other) tasks could the dataset be used for? There are many more tasks that can be framed for other challenges related to world modeling from this dataset. Some immediate examples: (1) offline reinforcement learning for game agents through imitation learning—predicting the sequence of actions that finish the game based on walkthroughs and reward information; (2) knowledge graph verbalization, a form of the standard data-to-text natural language processing task [Wiseman et al., 2017], in which we learn to generate text that is conditioned on a knowledge graph; and (3) description generation conditioned on the names of various objects, locations, and characters—with applications in long-form text generation domains such as automated storytelling [Martin et al., 2018, Fan et al., 2019] and procedural generation of interactive narratives [Ammanabrolu et al., 2020a, Walton et al., 2020].

Is there anything about the composition of the dataset or the way it was collected and preprocessed/cleaned/labeled that might impact future uses? For example, is there anything that a future user might need to know to avoid uses that could result in unfair treatment of individuals or groups (e.g., stereotyping, quality of service issues) or other undesirable harms (e.g., financial harms, legal risks)? If so, please provide a description. Is there anything a future user could do to mitigate these undesirable harms? Users should keep in mind that these come from games and can potentially describe non-normative situations.

Are there tasks for which the dataset should not be used? If so, please provide a description This dataset should not be used for tasks that involve direct physical interactions with humans, such as robotics.

B.6 Distribution

Will the dataset be distributed to third parties outside of the entity (e.g., company, institution, organization) on behalf of which the dataset was created? If so, please provide a description. It is open-sourced.

How will the dataset will be distributed (e.g., tarball on website, API, GitHub)? Does the dataset have a digital object identifier (DOI)? The dataset will be open-sourced at <https://github.com/JerichoWorld/JerichoWorld>.

When will the dataset be distributed? It was first released in May 2021.

Will the dataset be distributed under a copyright or other intellectual property (IP) license, and/or under applicable terms of use (ToU)? If so, please describe this license and/or ToU, and provide a link or other access point to, or otherwise reproduce, any relevant licensing terms or ToU, as well as any fees associated with these restrictions. The dataset will be under an MIT license, this is indicated on the GitHub repository.

Have any third parties imposed IP-based or other restrictions on the data associated with the instances? If so, please describe these restrictions, and provide a link or other access point to, or otherwise reproduce, any relevant licensing terms, as well as any fees associated with these restrictions. No.

Do any export controls or other regulatory restrictions apply to the dataset or to individual instances? If so, please describe these restrictions, and provide a link or other access point to, or otherwise reproduce, any supporting documentation. No.

B.7 Maintenance

Who is supporting/hosting/maintaining the dataset? Prithviraj Ammanabrolu will be responsible for maintenance.

How can the owner/curator/manager of the dataset be contacted (e.g., email address)? raj.ammanabrolu@gatech.edu or by filing an issue on the GitHub.

Is there an erratum? If so, please provide a link or other access point No.

Will the dataset be updated (e.g., to correct labeling errors, add new instances, delete instances)? If so, please describe how often, by whom, and how updates will be communicated to users (e.g., mailing list, GitHub)? Yes, more games will be added and corresponding data will be collected. Previous versions will be kept for backwards compatibility.

If the dataset relates to people, are there applicable limits on the retention of the data associated with the instances (e.g., were individuals in question told that their data would be retained for a fixed period of time and then deleted)? If so, please describe these limits and explain how they will be enforced. No.

Will older versions of the dataset continue to be supported/hosted/maintained? If so, please describe how. If not, please describe how its obsolescence will be communicated to users. Yes, versions will be archived on the GitHub repository.

If others want to extend/augment/build on/contribute to the dataset, is there a mechanism for them to do so? If so, please provide a description. Will these contributions be validated/verified? If so, please describe how. If not, why not? Is there a process for communicating/distributing these contributions to other users? If so, please provide a description They can fork and submit pull requests to the current repository if they wish to extend it—these will be validated in an open-source manner on GitHub via reviews of the extensions.

C Author Statement

The authors bear all responsibility in case of violation of rights, and confirm that this dataset is open-sourced under the MIT license.