Deep Dungeons and Dragons: Learning Character-Action Interactions from Role-Playing Game Transcripts

Annie Louis, Charles Sutton (2018)
https://aclanthology.org/N18-2111/

Presented by Tristan Galcik

Overview

1. Introduction
2. Dataset Preparation
3. Methodology (ACTION-LM, ACTION-LMS, CHAR-LM, CHAR-LMS)
4. Results
5. Paper Strengths/Weaknesses
6. Relation to Interactive Fiction & Story Generation

Introduction

Problem: how do we model interactions between events and characters?

- Use collaboratively-told RPG transcripts as a dataset for training neural language models
- Models aim to predict character-action interactions from text

Dataset Preparation

- Corpus made from 1,544 RPGs from roleplayerguild.com
  - ~56,000 posts, ~25 million tokens
- Each RPG has a character thread and gameplay thread
  - Character thread = character description posts
  - Gameplay thread = posts per player turn (action description)
- Story develops collaboratively due to each player's post in gameplay thread

Character description	**Action description**
Name: Ana Blackclaw; Age: 27; Gender: Female Appearance: Standing at a mighty 6'5, she is a giant among her fellow humans. Her face is light, though paler than the average man or woman's, and is marked by scars. ... Her body is muscular, as it would have to be to carry both her armor and the hammer. Her light grey eyes nearly always keep a bored expression. Her canines seem a tad larger than the normal person's. Preferred Weapon: Hammer. Preferred Armor: Heavy. Gift: Binoculars. Darksign: No.	She stopped dead in her tracks as the hissing began. A grumble escaped her as it did so, and she looked over to make sure the other woman was doing fine. Seeing that all was not entirely well, she allowed herself to slide down, her hand gripping the slope side once more to slow herself. Once that was accomplished, she reached out and grabbed the back of the girl's neck, pulling her back to steady herself. The giant remained silent as she did so, and then glanced over to the nearby skeletons. They would be upon them soon. Her grip tightened on the hammer as she glanced from side to side. It would not be a fun fight.

Dataset Preparation (continued)

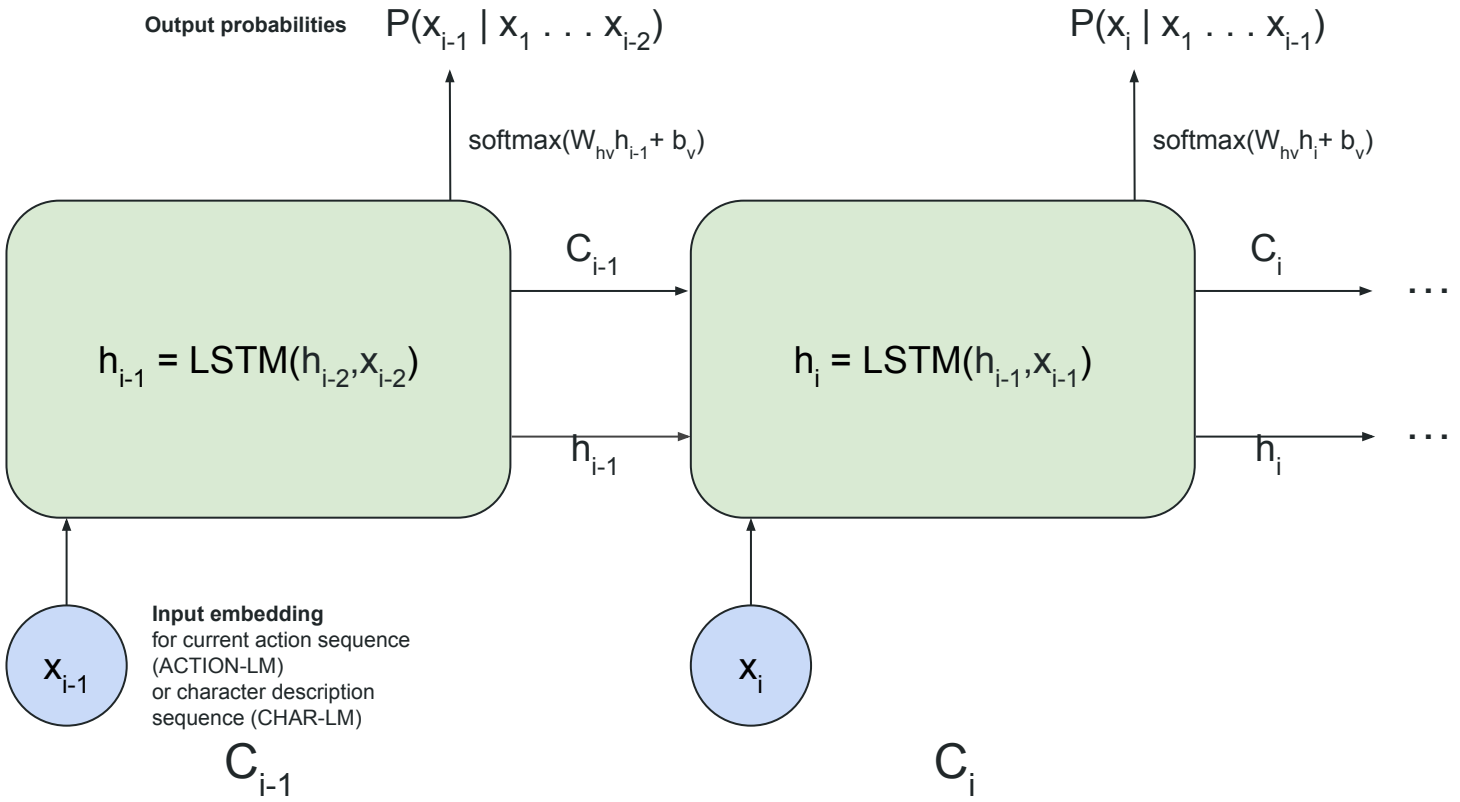
- Preprocess each character-action description to only keep parts relevant to the *character in focus*
  - Some contain descriptions/interactions of other characters (parents, allies, etc.)
- Only keep sentences with:
  - Character's name
  - Personal pronouns
  - Special tokens (personality, skill, ability, etc.)
- Replacement
  - Main character: "ENT" token
  - Other names: "NAME" token
  - Numbers: "NUM" token
  - Drop punctuation and sentences < 5 tokens
- Final dataset size:
  - 1,439 stories
  - 1.5 million action tokens
  - 2.95 million character tokens

Methodology

- Built separate models for action sequences and character descriptions
- Action sequence models
  - Baseline RNN model “**ACTION-LM**”, predicts next token in a sequence of action description
  - Character description augmented RNN model “**ACTION-LMS**”, each token in action sequence has information about the character incorporated
- Character description models
  - RNN model “**CHAR-LM**”, predicts next token for generating character descriptions
  - Action sequence augmented RNN model “**CHAR-LMS**”, conditioned on action sequence relevant to the character

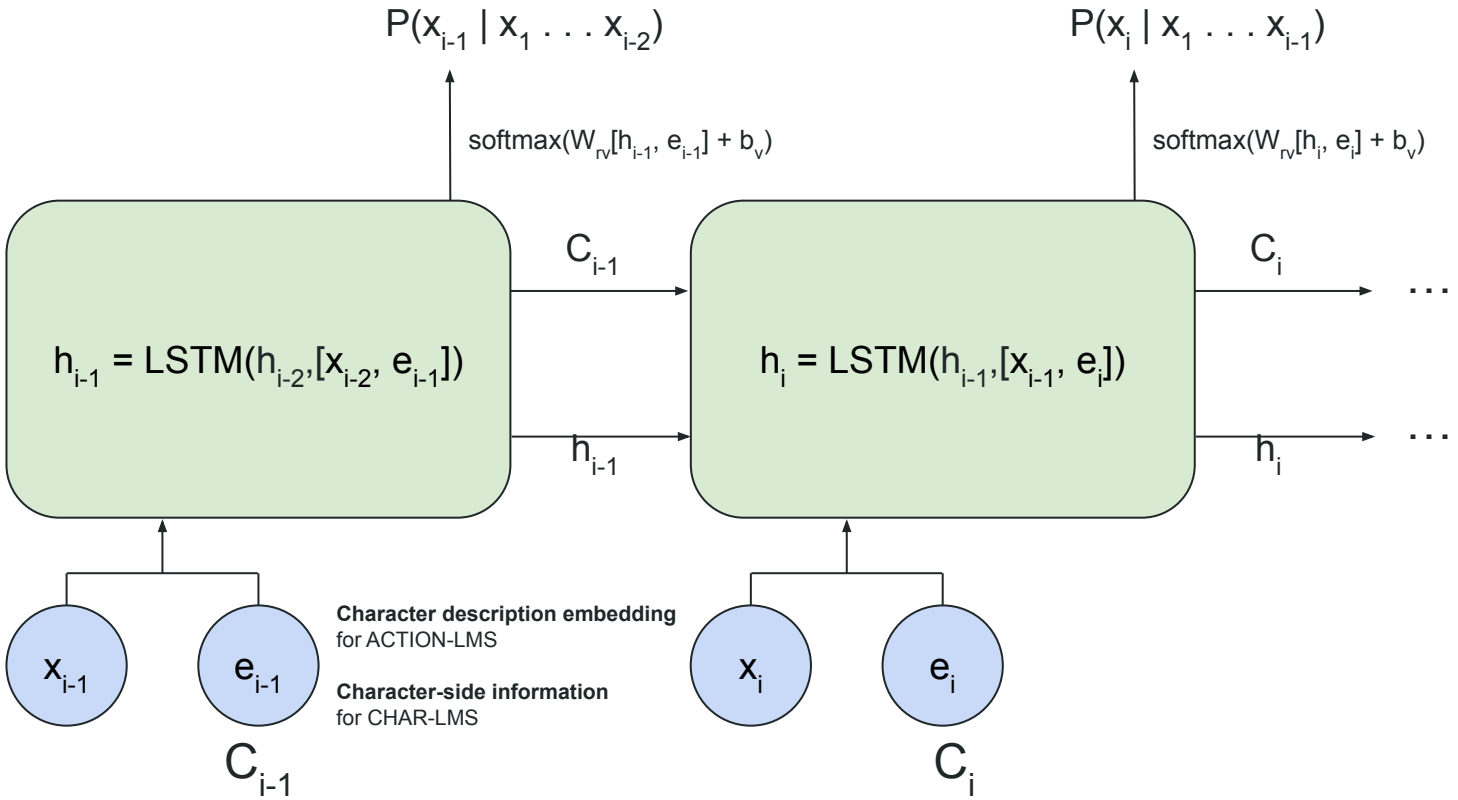
https://colah.github.io/posts/2015-08-Understanding-LSTMs/

ACTION-LM / CHAR-LM



https://colah.github.io/posts/2015-08-Understanding-LSTMs/

ACTION-LMS / CHAR-LMS



Results - Pattern Capturing

Generated continuations based on action/character context

ACTION-LMS

Char. context	Generated continuation
angry irritated	… back at name with her thick road with disappointment <eos>
hunter bow forest	… over and walked over to large king had been making sure

CHAR-LMS

Action context	Generated continuation
appeared disappeared flew	… a very young man who has a few scars on his body <eos>
waited	… a little girl who is a little girl who is a little

Results - Perplexity

Perplexity - measure of uncertainty, lower is better (https://mbernste.github.io/posts/perplexity/)

Model	Train	Dev	Test
ACTION-LM	82.56	106.83	105.06
ACTION-LMS	57.38	94.95	96.91
CHAR-LM	69.45	118.78	106.12
CHAR-LMS	61.84	110.13	100.86

Paper Strengths

- Dataset preparation (replacement) reduces noise and irrelevant information
- Use of baseline models for comparing results from ACTION-LMS and CHAR-LMS
- Relatively lightweight models, can be trained in ~1 hour

Paper Weaknesses

- Generated continuations are simplistic, limited in variety, and sometimes repetitive/incoherent
- Perplexity results show potential overfitting - train results are much lower (better) than dev/test results, indicates difficulty generalizing
- Could have used more methods of evaluating generated text
  - BLEU, ROUGE
  - Human evaluation

Relation to Story Generation and Interactive Fiction

- Could be used with NPCs - given an action sequence/character description context, generate dialogue
  - Example: given action context of a village under siege, NPC might seek help instead of sticking to generic behavior based on initial character description
- Procedural story generation - generated actions of characters drive the plot
- Co-creativity integration as a story writing assistant
  - Example: ACTION-LMS suggests action sequence continuations for stories
  - Example: CHAR-LMS suggests more details for a character description

Q&A