Assignment 10 script

Hello,

Today I am going to present a summary of a paper 'learning knowledge graph-based world models of textual environments' by prithviraj ammanabrolu and mark o riedl.

Text games are games where players interact with the world through textual natural language.

World models is abstract representation of the game world which improve an agent's ability to operate in interactive environments. They help predict changes in the world depending on one's actions which will help to plan future actions.

Previous models extracted information from surrounding while navigating new environments through rules, QA or transformer based extraction. These models benefit from memorization but lack the ability to predict change in graph state representation. During information extraction, relevant actions are overwhelmed with irrelevant actions. There are 2 challenges – what actions can I perform?' and 'how will the world change if I perform a particular action?'

This paper does 4 four contributions:

Represent changes in the world as difference between subsequent knowledge graph state representation

Introduce a worldformer which is a multi task transformer based architecture that learns the set of graph differences and relevant actions

It a introduces a loss function and training methodology to train the worldformer

Lastly it presents a study conducted on novel text games to show significance of above contributions

For model training, Jericho world dataset is used which has 24198 mapping of knowledge graph as tuple set (s, r, o) and natural language actions that cause changes in the world state. Each instance is (St, A, St+1, R). there are 2 tasks: knowledge graph generation and valid action generation

For knowledge graph generation, we predict difference between Gt and Gt+1. Following observations were made,

Gt+1 – Gt is a set of tuples added to Gt. Gt – Gt+1 is a set oof tuples deleted from Gt. Locations are fixed and unique in the text game. Objects and characters can only be in 1 location at a time and they cannot be open and closed simultaneously

To achieve knowledge graph generation, we have to predict the nodes to be added to Gt at time step t to transform into graph Gt+1 given text observations, valid actions, graph at time t and actions A for all the samples in the dataset

Worldformer takes as input textual observations, valid actions and graph at time step t.

Valid actions (Vt) are encoded to Ot and Graph at time step t (Gt) is encoded to Gt. Both are passed to aggregator to get combined encoded state St

There are 2 decoders in the worldformer. Action decoder and graph decoder. Internally there are designed the same way. They are conditioned on St directly and take Ot through cross-attention long with valid action Vt+1 and graph of next state Gt+1 as input respective

Decoding technique factors the distribution over target sequence into chain of conditional probabilities by using the given formula. Here theta is overall network parameters. Y is target sequence, X is input context via encoders

From this we calculate the maximum likelihood loss with cross-entropy at every step using the given formula.

We group all elements into Ysos (set of sequence form) for the set of sequence generation to train the worldformer using the given formula.

We do the same factorization distribution over output sos as chain of probabilities using the given formula. Ysos is independent of other elements so sequence of elements does not matter but tokens in each elements are dependent on each other so sequence of tokens in elements matters

We then calculate maximum likelihood loss for set of sequence to train the model using given formula

Combined loss from both decoders is given using this formula

Evaluation of the worldformer is done using 2 metrics: exact match and F1. At graph level true positive is when (s, r, o) triple all 3 matches with the ground truth triple. For token level, any relation or entitt must match the ground truth to be true positive. Positive exact match and F1 score happens only when all tokens in the predicted valid actions match the gold set

After evaluation, its observed that the worldformer performs than other baselines on the graph level and it outperforms the seq2seq model in valid actions prediction

In conclusion, we presented the worldformer for text games. Simplify the knowledge representation problem into predicting graph difference between states. Improved performance due to multi task training. Improved performance due to set of sequence loss instead of normal sequence loss.

Thankyou