Counting to Explore and Generalize in Text-based Games
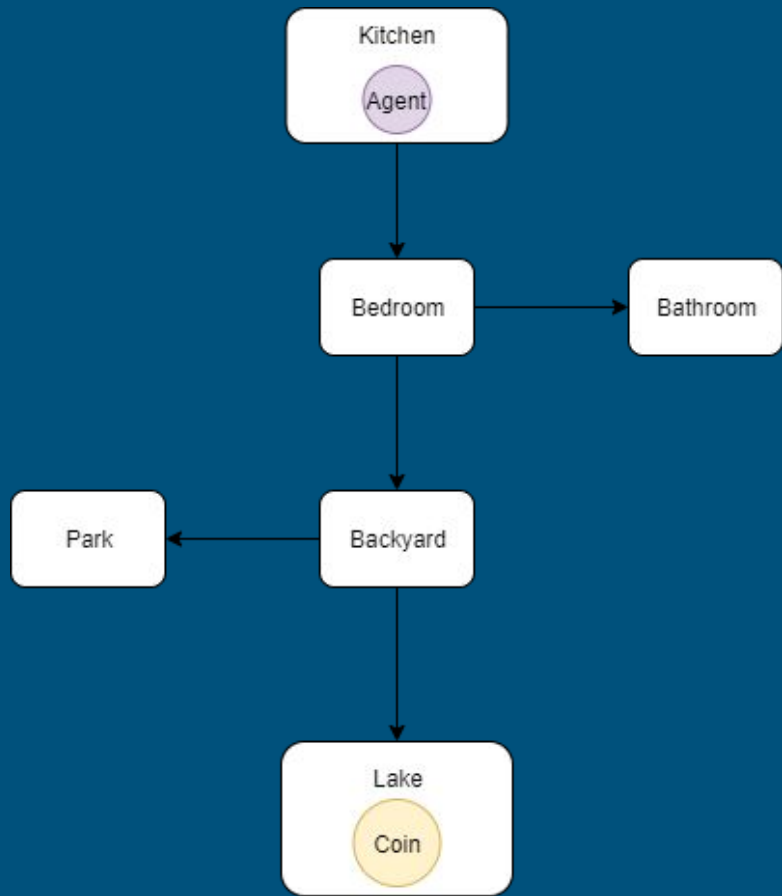
Xabier Benavides
and
David Sandstrom

Chain text game

- Nodes connected by paths.

- The paths create a chain.

- Distractor nodes that are off-chain.

- **Goal:** Find the coin in the end of the chain.

Explanation of the models

- Two different architectures: ***LSTM-DQN*** and ***LSTM-DRQN***.

- Two reward functions that promote exploration:

  - ***Cumulative counting bonus:*** Counts states since the <u>beginning of the training</u>.

  - ***Episodic discovery bonus:*** Counts states since the <u>beginning of the episode</u>.

Experiment

- While training, the <u>episodic discovery bonus</u> performed better.

- To simplify, we only tested the models trained with the episodic approach.

- Test were performed on <u>unseen games</u>.

Results

test lv 30, easy

DQN

test lv 30, easy
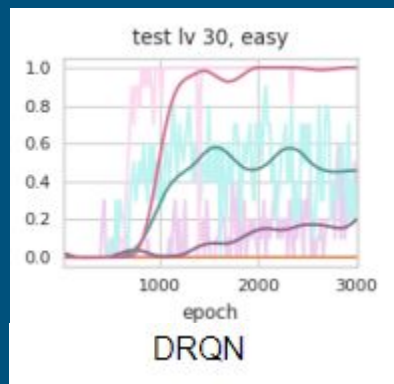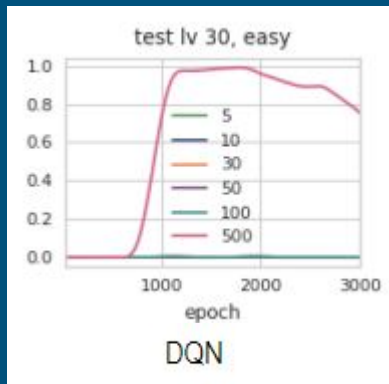
DRQN

- Easy games: DRQN model needs less training games.

- Hard games: Both models suffer from overfitting. DQN model learns a general strategy.

Experiences with the implementation

- Very complete.
  - World Generation.
  - LSTM-DQN.
  - LSTM-DRQN.

- Installation Problems.
- Unable to Run.

Conclusion

- Promising results in simple environments.

- *Generalization*.

- Future improvements in the field.