THEORY OF MIND NET

The Theory of mind network is aimed to understand machine learning scalability and reliability of its autonomous behaviour in an environment where different agents have, different “visibility” of the “maze”, the “agent” relies on a vector of action probabilities for its current state in order to “walk” through the maze, the bot itself ,this particular system is called a partially observed markov decision process POMDP (if the agent as full visibility then it is Markov decision process). This network can be used to make better approximations of multi agent tasks where the agents operate in an “observable” environment.

The network comprises of two different input layer, each with their own embedding layer,

- Character input layer is fed the past trajectories for each “episode” i.e. the iteration of the reinforcement learning for the agent. This was then passed onto an embedding layer, these embeddings are the sum of past input trajectories.
- The mental input layer is fed with just the current iteration trajectory up to the step before the current one, and the embedding layer from the character net

These two input layers are then fed into the prediction net which in the end outputs the values of its next step probabilities and predicted successor states i.e. estimating the agents policy, i.e the future steps.