## a. RL-LSTM

**ENVIRONMENT** 





 $(I_t, v_t, r_{t-1}, T_t) \longrightarrow o_t \longrightarrow e_t$ ENCODER  $e_t$ 











Policy Loss

