Simulated Robot Learning Interactively

$$P(future|state, action) = \frac{1}{P(state, action)} \langle P(future|history, action), P(action|history), P(state|history) \rangle_{P(history)}$$

$$P_{opt}(action|history) = \frac{P(action)}{Z_A(history, \mu)} e^{\frac{-1}{\mu} E_A(action, history)}$$

$$P_{opt}(state|history) = \frac{P(state)}{Z_S(history, \lambda)} e^{\frac{-1}{\lambda} E_S(state, history)}$$