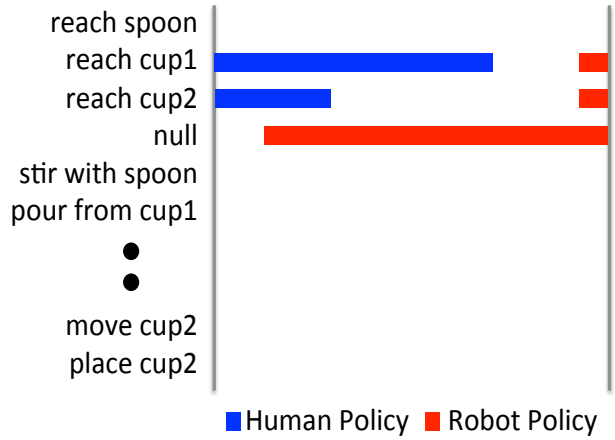
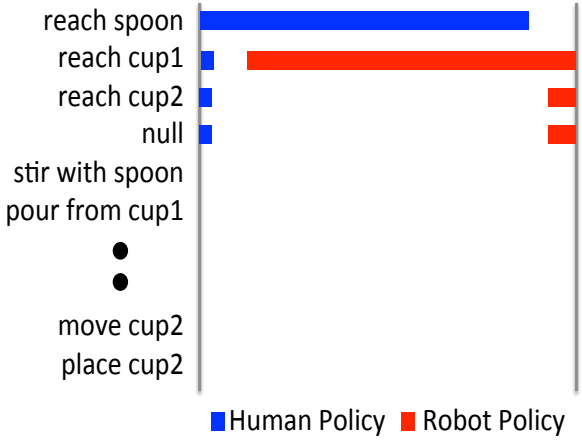
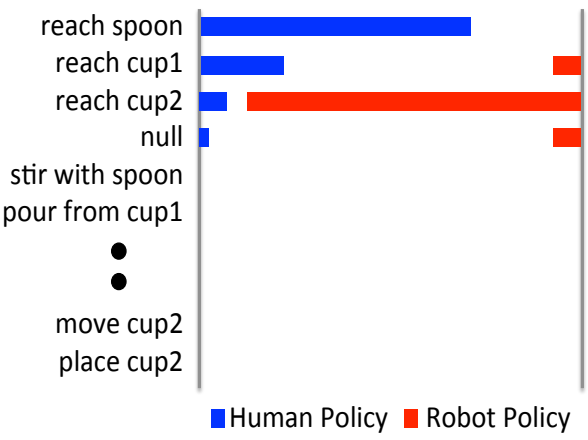


(a) Habit-following Human.



(b) ϵ-optimal Human.



(c) Adaptive Human.

Fig. 4: **Illustration of policies learnt with different human behaviors:** Each figure shows the learnt probability distributions of the various possible actions at the start state of the *following recipe* activity. Blue and red bars represent the probability of choosing the corresponding actions by the human and robot respectively.