Policy Comparison Metric

Policy comparison for rounds of a given game played between humans, and agents learning over given examples:

$S =$ set of total states in a game

$S_H =$ set of states seen by human agents

$E =$ set of states present/ seen in both agent and human games

$\eta_S = \dfrac{|E|}{|S_H|} =$ state similarity metric

$B_{s,a} =$ number of times action a is taken from state $s \in E$ by the agents

$B_s =$ most frequent action from s by agents $= \arg\max_a B_{s,a}$

$H_{s,a} =$ number of times action a is taken from state $s \in E$ by human players

$H_s =$ most frequent action from s by humans $= \arg\max_a H_{s,a}$

$\eta_P = \dfrac{\sum_{s \in E} \mathbb{1}(B(s) == H(s))}{|E|} =$ number of times the actions with highest frequencies from a state match between humans and agents