Apoorv CTF

Policy Functions for Certain q

Current Policy Function:

$$\frac{x^{\frac{3}{2}}}{\sqrt{5}}$$

Old Policy Function:

$$\sqrt{\frac{x}{7}}$$

Reference Policy Function:

$$\frac{x^{0.3}}{10}$$

Reward and Input Probability Lists

Reward Values	Input Probabilities
1.27	0.6
1.3	0.5
0.6	0.2
0.7	0.5

Table 1: Reward and Input Probability List

Hyperparameters

Parameter	Value
ϵ	0.5
β	0.2

Table 2: Epsilon and Beta Values