

2025 march 26

Wednesday, March 26, 2025 12:54

$$P(\text{choose } a)$$

$$= P(\text{explore}) \cdot P(a | \text{explore})$$

$$+ P(\text{exploit}) \cdot P(a | \text{exploit})$$

$a$	$P$
L	$\epsilon \cdot \frac{1}{4}$
R	$\epsilon \cdot \frac{1}{4}$
$\overset{\text{highest } Q}{\Rightarrow} U$	$(1 - \epsilon) + \epsilon \cdot \frac{1}{4}$
D	$\epsilon \cdot \frac{1}{4}$