1: **procedure** ϵ-Greedy Bernoulli Bandit Problem
2: Initialize the α, β-vectors, the success $\mathbf{S}$ and fail $\mathbf{F}$ vectors and action set $\mathcal{A}$
3: **for** $k \in 1$ to horizon **do** ▷ Play the game for horizon rounds
4: $a \leftarrow \texttt{nothing}$
5: **if** $\texttt{rand} \leq \epsilon_k$ **then** ▷ With threshold ϵ_k, explore
6: $a \leftarrow \texttt{uniform}(\mathcal{A})$ ▷ Exploration generates a uniform random $a \in \mathcal{A}$
7: **else** ▷ With threshold $1 - \epsilon_k$, exploit our current knowledge
8: $\mathbf{p} \leftarrow \{\text{Beta}(\alpha(a) + \mathbf{S}(a), \beta(a) + \mathbf{F}(a)) \mid \forall a \in \mathcal{A}\}$ ▷ Draw $|\mathcal{A}|$ samples
9: $a \leftarrow \arg\max_a \mathbf{p}$ ▷ Select action with the highest success probability
10: **end if**
11: $r \leftarrow \texttt{world}(a)$ ▷ Observe the reward $r \in \{0, 1\}$ from the `world`
12: $\mathbf{S}(a) \leftarrow \mathbf{S}(a) + r$ ▷ Update the success count for action a
13: $\mathbf{F}(a) \leftarrow \mathbf{F}(a) + (1 - r)$ ▷ Update the fail count for action a
14: **end for**
15: **end procedure**