Algorithm 1 CPSRL (Continuing Posterior Sampling for Reinforcement Learning)

1: **Input:** Prior dstribution f, discount factor γ, total earning time T
2: **Initialize:** $t = 1$, $k = 1$, $X_1 = 0$
3: **for** $t \leq T$ **do**
4:     **if** $X_t = 0$ **then**
5:         $t_k \rightarrow t$
6:         sample $\mathcal{E}^k \sim f(\cdot | \mathcal{H}_{t_k})$
7:         compute $\pi_k = \pi^{\mathcal{E}^k}$
8:         $k \leftarrow k + 1$
9:     **end if**
10:     sample and apply $A_t \sim \pi_k(\cdot | S_t)$
11:     observe R_{t+1} and S_{t+1}
12:     $t \leftarrow t + 1$
13:     sample $X_{t+1} \sim \text{Bernoulli}(\gamma)$
14: **end for**