Algorithm 1 BenchRL-QAS

Require: Set of quantum tasks $\mathcal{T}$ (e.g., VQE, VQSD, VQC, state preparation),
RL algorithms $\mathcal{A}$,
Encoding scheme (e.g., tensor-based),
Illegal action handler,
Curriculum schedule.

for each task t in $\mathcal{T}$ **do**
  Initialize environment $\mathcal{E}_t$ with Hamiltonian or objective for t
  for each RL algorithm a in $\mathcal{A}$ **do**
    Initialize agent $\mathcal{R}_a$ with action and state spaces
    Initialize quantum circuit C (an empty circuit)
    while not converged **do**
      Encode current circuit state $s \leftarrow \texttt{EncodeState}(C$
      Select action $u \leftarrow \mathcal{R}_a.\texttt{SelectAction}(s)$
      if $\texttt{IsIllegalAction}(u)$ **then**
        Apply penalty or mask action, update agent
        Continue to next iteration
      end if
      Update circuit $C \leftarrow \texttt{ApplyAction}(C, u)$
      Simulate circuit, obtain reward $r \leftarrow \mathcal{E}_t.\texttt{Evaluate}(C)$
      Observe new state $s' \leftarrow \texttt{EncodeState}(C)$
      $\mathcal{R}_a.\texttt{Update}(s, u, r, s')$
    end while
    Log performance metrics for (t, a)
  end for
  Optionally (if used curriculum learning): Increase task difficulty
end for
Aggregate and compare results across tasks and algorithms