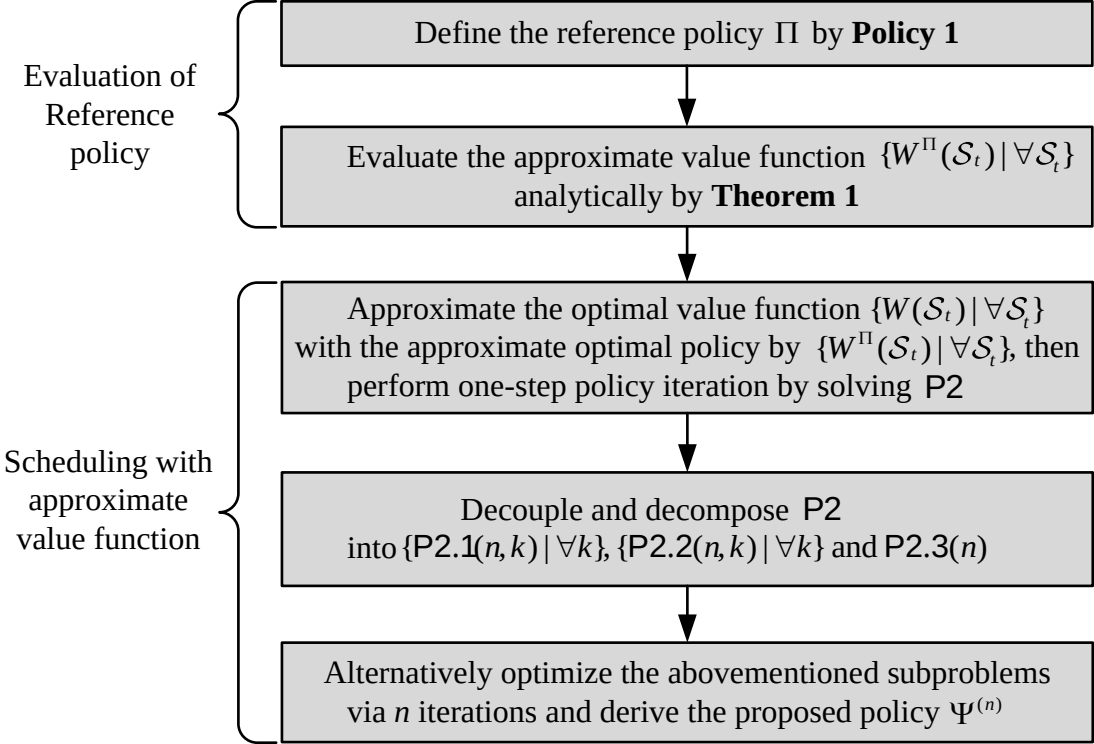
Evaluation of Reference policy
Define the reference policy Π by **Policy 1**
Evaluate the approximate value function $\{W^{\Pi}(\mathcal{S}_t) \mid \forall \mathcal{S}_t\}$ analytically by **Theorem 1**
Scheduling with approximate value function
Approximate the optimal value function $\{W(\mathcal{S}_t) \mid \forall \mathcal{S}_t\}$ with the approximate optimal policy by $\{W^{\Pi}(\mathcal{S}_t) \mid \forall \mathcal{S}_t\}$, then perform one-step policy iteration by solving P2
Decouple and decompose P2 into $\{\mathsf{P2.1}(n,k) \mid \forall k\}$, $\{\mathsf{P2.2}(n,k) \mid \forall k\}$ and $\mathsf{P2.3}(n)$
Alternatively optimize the abovementioned subproblems via n iterations and derive the proposed policy $\Psi^{(n)}$