function POMDP-VALUE-ITERATION($pomdp, \epsilon$) **returns** a utility function

inputs: $pomdp$, a POMDP with states S, actions $A(s)$, transition model $P(s' \mid s,a)$,
sensor model $P(e \mid s)$, rewards $R(s,a,s')$, discount γ
ϵ, the maximum error allowed in the utility of any state

local variables: U, U', sets of plans p with associated utility vectors α_p

$U' \leftarrow$ a set containing all one-step plans $[a]$, with $\alpha_{[a]}(s) = \sum_{s'} P(s' \mid s,a)\, R(s,a,s')$

repeat

$U \leftarrow U'$

$U' \leftarrow$ the set of all plans consisting of an action and, for each possible next percept, a plan in U with utility vectors computed according to Equation (16.18)

$U' \leftarrow$ REMOVE-DOMINATED-PLANS(U')

until MAX-DIFFERENCE(U, U') $\leq \epsilon(1-\gamma)/\gamma$

return U