1. function $value\_iteration(S, A, p, r, \lambda)$ return optimal policy v^*
2. initialise v^* randomly
3. repeat
4. $v := v^*$
5. for each $s \in S$ do
6. $$v^*(s) := \max_{a \in A} \left(r(a,s) + \lambda \sum_{s' \in S} p(s' \mid s, a) v(s') \right)$$
7. end-for
8. until v and v^* are sufficiently close
9. return v^*
10. end function $value\_iteration$