Some parts of the present work were initiated in [3], and developed in a different direction in [4, 12]. It is remarkable that Howard'

where $x \in (\mathbb{R}_{\max})^n \setminus \{\mathbb{0}\}$

policy, which is a map

$$\pi : \mathcal{N} \to \mathcal{E}, \quad \text{such that} \quad \text{In}(\pi(i)) = i, \ \forall i \in \mathcal{N} \ .$$

That is, a policy is just a map which with a node

2. *Policy improvement*. Improve the policy π , using Algorithm IV.3 with input $\pi = \pi$, $\eta = \eta$