evaluation

$Q \rightsquigarrow q_\pi$

π

Q

$\pi \rightsquigarrow \text{greedy}(Q)$

improvement