
Algorithm 1 Q-Learning

initialize Q-table

initialize S

$f_{\text{fatigue}} \leftarrow 0$

for athlete i : **do**

 randomly initialize the parameter vector $\vec{v} = (v_1, v_2, \dots, v_k)$ of i # showing his parameters

 his current strategy vector $v(\vec{t}) = (v_1(t), v_2(t), \dots, v_k(t))$ # showing his strategy

end for

while not at the ending point: **do**

 # change his strategy vector $v(\vec{t})$ according to Q-table

 choose a strategy A based on $v(\vec{t})$ and P

 calculate the expected (best) strategy for the current terrain based on previous calculations $\vec{v}^\Theta = (v_1^\Theta, v_2^\Theta, \dots, v_k^\Theta)$

$$\Delta^\Theta v \leftarrow \sqrt{\frac{\sum_{i=1}^k (v_i - v_i^\Theta)^2}{k}} \stackrel{\text{def}}{=} f_\Delta(\vec{v})$$

$$Q(S, A) = (1 - \underbrace{\alpha}_{\text{learning rate}})Q(S, A) + \underbrace{\alpha}_{\text{learning rate}} [\underbrace{R}_{\text{reward}} + \underbrace{\gamma}_{\text{discount factor}} \cdot \text{inverse strategy}(\nabla f_\Delta(\vec{v})) -$$

$Q(S, A)]$ # learning rate α can change due to f_{fatigue}

 # when f_{fatigue} increases, α decreases

 Update Q-table

 Update f_{fatigue}

 Update α

end while

show the best strategy
