

состояние $f_1(x_t)$

Агент
Q-learning

выигрыш

$$f_1(x_t) - f_1(x_{t-1})$$

функция
приспособленности
(f_0 или f_1)

(1 + 1)-EA

