Иногда стреляют вправо сразу и это правильно

n\_episodes = 30 # количество эпизодов lapan = 20

ПОПЫТКА ПОВТОРНО ОБУЧИТЬ на следующий день НЕУДАЧА!!!!

alpha = 0.9 #learning rate sayon - 0.5 больш - 0.9 Lapan = 0.2

gamma = 0.9 #discount factor sayon - 0.9 больш - 0.5 lapan = 0.9

epsilon = 0.1 #e-greedy sayon - 0.3 больш - 0.7 lapan = = 1.0 (100% random actions)

n\_statesFox = 16 # количество состояний нашего мира-сетки

get\_stats()= {'battles\_won': 3, 'battles\_game': 10, 'battles\_draw': 4, 'win\_rate': 0.3, 'timeouts': 4, 'restarts': 0}

actions\_history= [2, 3, 2, 3, 3, 3, 4, 3, 4, 3, 4, 3, 4, 3, 4, 3, 4, 3, 5, 2, 4, 3, 5, 2, 4, 3, 5, 2, 4, 3, 5, 2, 4, 3, 5, 2, 4, 3, 5, 2, 3, 3, 4, 2, 5, 3, 4, 2, 5, 3, 4, 2, 5, 3, 5, 2, 4, 3, 5, 2, 4, 3, 5, 2, 4, 3, 5, 3, 5, 2, 4, 3, 5, 2, 4, 3, 5, 2, 5, 3, 4, 2, 5, 3, 4, 3, 5, 2, 6, 3, 5, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2]

Average reward = 8.037681842672416

get\_stats()= {'battles\_won': 6, 'battles\_game': 30, 'battles\_draw': 11, 'win\_rate': 0.2, 'timeouts': 11, 'restarts': 0}

actions\_history= [2, 3, 2, 3, 3, 3, 4, 3, 4, 3, 4, 3, 4, 3, 4, 3, 4, 3, 5, 2, 4, 3, 5, 2, 4, 3, 5, 2, 4, 3, 5, 2, 4, 3, 5, 2, 4, 3, 5, 2, 4, 3, 5, 2, 4, 3, 5, 2, 3, 3, 5, 2, 4, 3, 5, 2, 4, 3, 5, 2, 4, 3, 5, 2, 5, 3, 4, 3, 5, 2, 4, 3, 5, 2, 4, 3, 5, 2, 6, 3, 6, 2, 6, 3, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 6, 3, 6, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 3, 0, 2, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2]

Average reward = 6.247196030890806

ПОПЫТКА ПОВТОРНО ОБУЧИТЬ на следующий день НЕУДАЧА!!!!

get\_stats()= {'battles\_won': 0, 'battles\_game': 10, 'battles\_draw': 0, 'win\_rate': 0.0, 'timeouts': 0, 'restarts': 0}

actions\_history= [3, 3, 3, 3, 3, 3, 4, 3, 4, 3, 4, 3, 4, 3, 4, 3, 4, 3, 4, 2, 4, 3, 4, 2, 4, 3, 4, 2, 4, 3, 4, 2, 4, 3, 4, 2, 4, 3, 4, 2, 4, 3, 4, 2, 4, 3, 4, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 7, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3, 0, 2, 0, 3]

Average reward = 3.2906115301724164

ПОПЫТКА 2 ПОВТОРНО ОБУЧИТЬ на следующий день НЕУДАЧА!!!!

get\_stats()= {'battles\_won': 0, 'battles\_game': 30, 'battles\_draw': 0, 'win\_rate': 0.0, 'timeouts': 0, 'restarts': 0}

actions\_history= [3, 2, 3, 2, 3, 4, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 4, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1]

Average reward = 2.478007139008626

ПОПЫТКА 3 ПОВТОРНО ОБУЧИТЬ на следующий день НЕУДАЧА НИЧЕГО НЕ ВЫУЧИЛИ ВООБЩЕ

get\_stats()= {'battles\_won': 0, 'battles\_game': 30, 'battles\_draw': 0, 'win\_rate': 0.0, 'timeouts': 0, 'restarts': 0}

actions\_history= [3, 1, 3, 1, 3, 1, 2, 1, 5, 1, 3, 1, 2, 1, 3, 1, 2, 1, 5, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 5, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 5, 1, 3, 1, 3, 1, 2, 1, 5, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 5, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 5, 1, 3, 1, 2, 1, 3, 1, 2, 1, 5, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 5, 1, 3, 1, 3, 1, 2, 1, 3, 1, 2, 1, 5, 1, 3, 1, 2, 1, 3, 1, 2, 1, 5, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 5, 1, 3, 1, 2, 1, 3, 1, 2, 1, 5, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 5, 1, 3, 1, 3, 1, 2, 1, 3, 1, 2, 1, 5, 1, 3, 1, 2, 1, 3, 1, 2, 1, 5, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 3, 1, 2, 1, 2, 1, 3, 1, 3, 1, 2, 1, 2, 1, 3, 1, 2, 1, 3, 1, 3, 1, 2, 1, 3, 1, 2, 1, 2, 1, 3, 1, 3, 1, 2, 1, 2, 1, 3, 1, 2, 1, 3, 1, 3, 1, 2, 1, 3, 1, 2, 1, 2, 1, 3, 1, 2, 1, 3, 1, 0, 1, 0, 1, 0, 1]

Average reward = 0.03088631465517252