Задачи

1. Среда – клеточный мир (рис 1). Агент находиться в клетке 1, терминальное состояние 5-я клетка. Найти оптимальную стратегию методом Монте Карло, γ=1. В нулевом приближении стратегия случайная, равномерная (т.е. равновероятное движение агента влево и вправо).
2. Среда – клеточный мир (рис 1). Агент находиться в клетке 1, терминальное состояние 5-я клетка. Найти оптимальную стратегию методом временных различий, γ=1.
3. Среда – клеточный мир (рис 1). Агент находиться в клетке 1, терминальное состояние 5-я клетка. Найти оптимальную стратегию методом SARSA, γ=1.
4. Среда – клеточный мир (рис 1). Агент находиться в клетке 1, терминальное состояние 5-я клетка. Найти оптимальную стратегию методом Q-learning, γ=1.
5. Среда – клеточный мир (рис 1). Агент находиться в клетке 1, терминальное состояние 5-я клетка. Найти оптимальную стратегию методом crossentropy. Реализовать мнемонический код
6. Среда – клеточный мир (рис 1). Агент находиться в клетке 1, терминальное состояние 5-я клетка. Реализовать мнемонический код для нахождения оптимальной стратегии методом deep Q learning. Предложить архитектуру сети
7. Среда – клеточный мир (рис 1). Агент находиться в клетке 1, терминальное состояние 5-я клетка. Реализовать мнемонический код для нахождения оптимальной стратегии методом reinforce. Предложить архитектуру сети
8. Среда – клеточный мир (рис 1). Агент находиться в клетке 1, терминальное состояние 5-я клетка. Реализовать мнемонический код для нахождения оптимальной стратегии методом actor-critic. Предложить архитектуру сети
9. Многорукий бандит. Среда вернула последовательность наград для одного автомата: 0, 0, 1. Оценить значение верхней доверительной границы в методе USB1
10. Многорукий бандит. Для первого автомата среда вернула последовательность наград 0, 0, 1, для второго автомата последовательность 0, 1, 1, 0, 0. Выбрать руку на следующем шаге в соостветствии с эпсилон жадной стратегией и оптимизмом в неопределенности.
11. Кинули монетку 3 раза, результат Орел, Решко, Орел. Найти вероятность выпадения Орла методом максимального правдоподобия а также оценить распределение матожидания выпадения Орла, т.е. оценить апостериорное распределение не используя сопряженное распределение
12. Кинули монетку 3 раза, результат Орел, Решко, Орел. Найти вероятность выпадения Орла методом максимального правдоподобия а также оценить распределение матожидания выпадения Орла, т.е. оценить апостериорное распределение с использованием сопряженного распределения
13. Многорукий бандит. Для первого автомата среда вернула последовательность наград 0, 0, 1, для второго автомата последовательность 0, 1, 1, 0, 0. Выбрать руку на следующем шаге в соостветствии с методом UCB1
14. Многорукий бандит. Для первого автомата среда вернула последовательность наград 0, 0, 1, для второго автомата последовательность 0, 1, 1, 0, 0. Выбрать руку на следующем шаге в соостветствии с методом Томпсон семплинг.
15. Дана функция распределения sin(x)/2 на отрезке [0, pi]. Написать мнемонический алгоритм генерации случайных чисел на основе projection sampling.
16. Дана функция распределения sin(x)/2 на отрезке [0, pi]. Написать мнемонический алгоритм генерации случайных чисел на основе importence sampling.
17. Дана функция распределения sin(x)/2 на отрезке [0, pi]. Написать мнемонический алгоритм генерации случайных чисел используя схему Метраполиса - Гастингса



Рис 1.