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. 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 и сделать обоснование выбора.
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 не используя сопряженное распределение
13. 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
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 UCB1
15. 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.
16. Дана функция распределения sin(x)/2 на отрезке [0, pi]. Написать мнемонический алгоритм генерации случайных чисел на основе projection sampling.
17. Дана функция распределения sin(x)/2 на отрезке [0, pi]. Написать мнемонический алгоритм генерации случайных чисел на основе importence sampling.
18. Дана функция распределения sin(x)/2 на отрезке [0, pi]. Написать мнемонический алгоритм генерации случайных чисел используя схему Метраполиса - Гастингса
19. Данные приходят из нормального распределения с неизвестными параметрами. Модель данных y = wx + z. Здесь x – наблюдения, z – модель шума z~N(0,s2). Сделать приближенный вариационный вывод параметра w
20. Данные приходят из распределения Пуассона. Наблюдаемые величины x = [12,10,5,11]. В качестве априорного распределения на параметр λ (в распределении Пуассона) задается логнормальное распределение с парамерами μ и σ. Предложить схему уточнения λ на основе поступающей информации (данных xi)
21. Данные приходят из логнормального распределения с неизвестными парамерами μ и σ. Наблюдаемые величины x = [1.22,1.43,2.2]. Предложить априорные распределения на параметры μ и σ а также схему уточнения этих параметров на основе поступающей информации (данных xi)
22. Данные приходят из экспоненциального распределения. Наблюдаемые величины x = [1.2,2.1,1.7]. В качестве априорного распределения на параметр λ (в экспоненциальном распределении) задается логнормальное распределение с парамерами μ и σ. Предложить схему уточнения λ на основе поступающей информации (данных xi)
23. Данные приходят из геометрического распределения с параметром p. Наблюдаемые величины x = [5, 8, 6]. Предложить априорное распределение на параметр p и схему уточнения параметра на основе поступающей информации (данных xi)
24. Данные приходят из нормального распределения с параметрами μ и σ. Параметр μ нам известен μ=1. Наблюдаемые величины x = [0.9, 1.04, 1.22]. Предложить априорное распределение на параметр σ и схему уточнения параметра на основе поступающей информации (данных xi)
25. Данные приходят из нормального распределения с параметрами μ и σ. Параметр σ нам известен σ=1. Наблюдаемые величины x = [0.9, 2.04, 1.22]. Предложить априорное распределение на параметр μ и схему уточнения параметра на основе поступающей информации (данных xi)



Рис 1.