Экзаменационные вопросы

RL

1. Строение RL агента. Стратегия, полезность, модель

2. Строение RL агента. Model free vs model base подходы

3. Марковские процессы принятия решений и их связь с обучение с подкреплением. Примеры.

4. Уравнение Белмана. Уравнение оптимальности Беллмана

5. Итерационные алгоритмы поиска оптимальной стратегии. Итерации по стратегии и итерации по полезностям.

6. Алгоритмы основанные на полезности. Алгоритм Монте Карло для поиска оптимальной стратегии. Достоинства и недостатки

7. Алгоритмы основанные на полезности. TD алгоритм. Достоинства и недостатки

8. Алгоритмы основанные на полезности. Q-learning. Достоинства и недостатки

9. Алгоритмы основанные на полезности. SARSA. Достоинства и недостатки

10. Аппроксимация функции полезности. Линейный подход. Целевая функция (функционал ошибки)

11. Аппроксимация функции полезности. Нелинейный подход. Deep Q Learning. Experience replay. Функционал ошибки

12. Аппроксимация функции полезности. Нелинейный подход. Double Q-learning. Dueling Network. Функционал ошибки

13. Алгоритмы основанные на стратегии. Crossentropy метод. Привести пример на основе клеточного мира. Функционал ошибки.

14\*. Алгоритмы основанные на стратегии. Policy gradient. Вывод алгоритма Reinforce.

15\*. Advantage actor-critic. Вывод.

Многорукий бандит и байесовская статистика

16. Многорукий бандит. Постановка задачи. Exporation vs exploitation. Regret и методы его оценки

17. Многорукий бандит. Стратегии исследований. Теорема Lia and Robbins

18. Многорукий бандит. Верхняя доверительная граница (UCB1). Неравенство Хёфдинга

19. Многорукий бандит. Томпсон семплинг. Пример с рекламой (монеткой) с формулами пересчета

20. Контекстный многорукий бандит. Подходы к решению: линейная байесовская регрессия vs нейросеть (байесовская нейросеть)

21. Формула Байеса. Подходы для расчета апостериорного распределения. Сопряженное априорное распределение. Примеры

22. Методы оценки (апостериорных) распределений. Projection sampling, importence sampling

23. Мonte Carlo Markov Chain. Марковские цепи. Стационарность и эргодичность. Алгоритм Метрополиса-Гастингса

24\*. Вариационные методы. EM алгоритм. Пример

25\*. Вариационные методы. Приближенный вариационный вывод

26\*. Вариационные методы. Mean field variational inference. Пример