

План курса

Продвинутые Темы в Обучении с Подкреплением

В интернете достаточно ресурсов и открытых курсов предоставляющих хорошее введение в методы Обучения с Подкреплением (Reinforcement Learning). Курс покрывающий введение в Обучение с Подкреплением сейчас читается в МФТИ. Идея этого курса заключается в том чтобы сконцентрироваться на современном направлении исследований в RL и разобрать значимые статьи за последние несколько лет.

Под каждой лекцией приведен краткий список обсуждаемых тем и ссылки на статьи которые предполагается обсудить в лекции.

Список лекций:

1. Многозадачное обучение и перенос знаний

- Пакетное многозадачное обучение[1]
- Обобщение стратегии агента на более сложные версии среды[2, 3]
- Перенос стратегии на новые цели[4]
- Sim2Real: Перенос стратегии из симулятора в реальный мир[5, 6]

2. Иерархическое обучение с подкреплением

- Иерархическая декомпозиция задачи
- Иерархическое обучение с заданными подцелями[7]
- Автоматическое выделение подцелей[8, 9]

3. Мета-обучение

- Априорная смещенность алгоритма обучения
- Мета-обучение с подкреплением[10, 11, 12]
- Непрерывное обучение и проблема “катастрофического забывания”[13]

4. Эпизодическая память агента и ЧНМППР

- Частично Наблюдаемый Марковский Процесс Принятия Решений (ЧНМППР)
- Рекуррентные сети в алгоритмах обучения с подкреплением: R2R2 [14]
- Нейросети дополненные механизмом памяти [15]
- Полу-параметрические архитектуры памяти[16]

5. Обучение на основе модели среды

- Планирование при наличии модели среды[17]
- Обучение модели среды[18]
- Оценка “неуверенности” модели среды[19]
- Комбинирование подходов обучения с явной моделью среды и без нее[20]

6. Подходы к исследованию среды

- Компромисс между “исследованием” и “использованием” во время обучения
- Исследование среды на основе подсчета состояний[21]
- Исследование на основе “внутренней мотивации”[22, 23, 24]

7. Распределенное обучение с подкреплением

- Оценка распределения значений функции ценности состояний[25]
- Квантильная Регрессия для распределенного обучения с подкреплением[26, 27]
- Анализ различий между поведением алгоритмов распределенного ОП и классического ОП[28]

8. Обучение на основе демонстраций

- Имитация стратегии на основе демонстраций
- Алгоритмы имитационного обучения: GAIL[29], DQfD[30]
- Обратное обучение с подкреплением[31, 32]

9. Генетические алгоритмы

- Генетические алгоритмы для обучения агента[33, 34]
- Композиция генетических алгоритмов с ОП [35, 36]
- Улучшение исследования среды через эволюционные стратегии[37]

10. Мультиагентное обучение с подкреплением

- Меж-агентный перенос знаний[38]
- Кооперативные и Соревновательные многоагентные стратегии[39, 40]
- Обучение меж-агентной коммуникации[41]

11. Масштабирование алгоритмов обучения с подкреплением

- Масштабирование алгоритмов с единой стратегией: IMPALA[42]
- Масштабирование алгоритмов с разделенной стратегией: Ape-X[43], R2D2[44]
- Открытые библиотеки алгоритмов обучения с подкреплением

12. Байесовские методы в обучении с подкреплением

- Soft Actor-Critic[45, 46]
- TD-обучение вариационного авто-энкодера [47]

13. Обучение с подкреплением для задач комбинаторной оптимизации

- Зачем использовать ОП для комбинаторной оптимизации?
- Использование архитектуры “Трансформер” для кодирования графовых структур [48]
- Интерпретируемая комбинаторная оптимизация с использованием алгоритмов ОП[49]
- Масштабирование на задачи большей размерности[50]