

شماره درس: 28656

نام درس: سیستم های مبتنی بر یادگیری تقویتی

نوع و بیشینه واحد: تخصصی اختیاری، ۳ واحد
همینا:
اولین نیمسال ارائه: نیم سال اول ۱۴۰۱-۱۴۰۰
آخرین ویرایش: تابستان ۱۴۰۰

نوع درس: نظری
پیشنیاز:
مقطع: تحصیلات تکمیلی
گروه:

اهداف: این درس به منظور آشنایی دانشجویان مقطع تحصیلات تکمیلی با طراحی سیستم های هوشمند مبتنی بر یادگیری تقویتی است، که در آن یادگیری سیستم بکمک تعامل عامل هایش با محیط انجام می شود. این درس در بر دارنده همه پیشنیاز های لازم برای دانشجویان مکانیک است. در حین اینکه دانشجویان در طی درس آموزش می بینند تا یک پروژه تحقیقاتی را بوسیله Python انجام دهند. این درس کاربرد زیادی در بسیاری از حوزه ها از قبیل رباتیک، کنترل، بیومکانیک، هوا فضا، روش های درمانی و تصمیم سازی در سیستم های پیچیده دارد

سرفصلها:

۱- مروری کلی بر سیستم های مبتنی بر یادگیری

۱-۱- معرفی انواع سیستم های مبتنی بر یادگیری

۱-۲- بخش های اساسی یک سیستم یادگیری تقویتی

۱-۳- نمونه هایی از سیستم های یادگیری تقویتی

۲- فرم رسمی مساله یادگیری تقویتی

۲-۱- انواع ساختارهای عامل های هوشمند

۲-۲- مشخصات محیط

۲-۳- ویژگی تصمیم گیری مارکوف در مسایل یادگیری تقویتی

۲-۴- توابع ارزش

۲-۵- سیاست بهینه و معادلات بهینگی بلمن

۳- روش های انتخاب تابع عمل

۳-۱- تخمین ارزش عمل

۳-۲- روش های جستجوی مستقیم و غیر مستقیم

۳-۳- بهبود روش جستجوی بیشینه نرم

۳-۴- روش انتخاب عمل، کران اطمینان بالا

۴- برنامه سازی پویا

۴-۱- ارزیابی سیاست

۴-۲- بهبود سیاست

۴-۳- تکرار سیاست

۴-۴- تکرار ارزش

۴-۵- کارآرایی برنامه سازی پویا

۵- روش های مونت کارلو

۵-۱- تخمین تابع ارزش با مونت کارلو

۵-۲- تخمین تابع ارزش حالت- عمل با مونت کارلو

۵-۳- کنترل مونت کارلو

۵-۴- کنترل مونت کارلو به روش بر سیاست

۵-۵- تکرار سیاست تعمیم یافته

۵-۶- کنترل مونت کارلو به روش برون سیاست

۶- یادگیری تفاضل- موقتی

۶-۱- پیشگویی تفاضل- موقتی

۶-۲- به روز رسانی دسته ای در روش یادگیری تفاضل- موقتی

۶-۳- کنترل در روش یادگیری تفاضل- موقتی

۶-۳-۱- الگوریتم یادگیری Sarsa

۶-۳-۲- الگوریتم یادگیری Q

۶-۳-۳- الگوریتم یادگیری Sarsa مورد انتظار

۶-۴- روش های ارزیابی الگوریتم های یادگیری تقویتی

۷- یادگیری تفاضلی موقتی n مرحله ای و شایستگی مسیر

۷-۱- روش های خود راه انداز n مرحله ای

۷-۱-۱- روش پیشگویی n مرحله ای تفاضل موقتی

۷-۱-۲- روش n مرحله ای Sarsa

۷-۲- شایستگی مسیر

۷-۲-۱- الگوریتم پیشگویی $TD(\lambda)$

۷-۲-۲- الگوریتم یادگیری $Sarsa(\lambda)$

۸- مباحث پیشرفته در یادگیری تقویتی

۸-۱- تعمیم در یادگیری تقویتی

۸-۲- یادگیری تقویتی عمیق و DQN

۸-۳- روش گرادیان سیاست

۸-۴- روش نقاد- عملگر

۵-۸- یادگیری تقلیدی

۸-۶- یادگیری تقویتی بیشینه آنتروپی

۹- ساختارهای مختلف شبکه های نقاد-عملگر برای کاربرد های کنترلی

۹-۱- روش گرادیان سیاست عمیق (DDPG)

۹-۲- روش گرادیان سیاست قطعی عمیق جفت تاخیری (TD3)

۹-۳- روش بازیگر - نقاد نرم (SAC)

مراجع:

- 1- Richard S. Sutton, Andrew G. Barto, “Reinforcement Learning: An Introduction”, second edition, MIT press, 2018.
- 2- Hao Dong, Zihan Ding, Shanghang Zhang, “Deep Reinforcement Learning, Fundamentals, Research and Applications, Springer, 2020.
- 3- Maxim Lapan, “Deep Reinforcement Learning Hands-On”, Second edition, Packt publishing, 2020.
- 4- Sudharsan Ravichandiran, “Hands on Reinforcement Learning with Python”, Packt publishing, 2018.
- 5- Andrea Lonza, “Reinforcement Learning Algorithms with Python”, Packt publishing, 2019.
- 6- Farrukh Akhtar, “Practical Reinforcement Learning, Develop Self Evolving, Intelligent Agents with Open AI Gym, Python and Java”, Packt publishing, 2017.
- 7- Mohit Sewak, “Deep Reinforcement Learning, Frontiers of Artificial intelligence”, Springer, 2019.