



دانشکده مهندسی
کامپیوتر و فناوری اطلاعات



دانشگاه صنعتی امیرکبیر
(پلی تکنیک تهران)

دانشگاه صنعتی امیرکبیر
دانشکده مهندسی کامپیوتر و فناوری اطلاعات

یادگیری تقویتی در روباتیک

یاسمن سادات میرمحمد

۱۰ آذر ۱۳۹۶

استاد راهنما:

دکتر رضا صفابخش

تکلیف شماره ۴

پیشنهاد پروژه کارشناسی

پاییز ۹۶

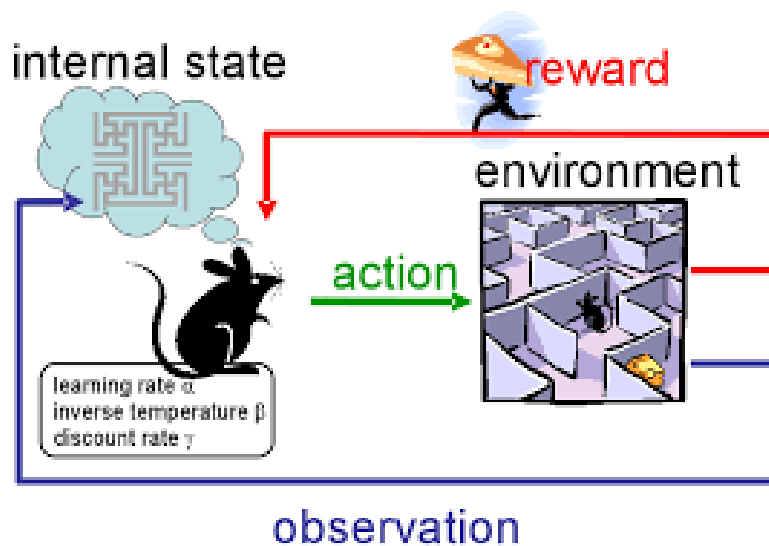
مقدمه

یادگیری تقویتی یکی از زیرمجموعه های علم یادگیری ماشین است. در این نوع یادگیری، یک سری عامل هوشمند تعریف میکنیم، وظیفه این عاملان این است که از طریق تعامل با محیط اطرافشان، مشاهده ی نتایج و چک کردن یک تابع پاداش و جزا(مثبت یا منفی) رفتار خود را به یک حالت بهینه نزدیک کنند.

این نوع یادگیری به یادگیری موجودات زنده و به خصوص انسان ها نزدیک است و از الگوهای مشابه پیروی میکند. ازین جهت میتواند برای اهداف خاصی در زمینه علم رباتیک، مانند ساخت ربات های انسان نما، ساخت هوش های مصنوعی که نیاز است رفتار انسانی داشته باشند و ... استفاده شوند. در واقع یادگیری تقویتی، از روانشناسی رفتارگرایی الهام میگیرد.

تفاوت اصلی میان یادگیری نظارت شده و یادگیری تقویتی این است که در یادگیری تقویتی، هیچ گاه به عامل گفته نمی شود که عمل صحیح در هر وضعیت چیست، و فقط به وسیله ی معیاری، به عامل گفته می شود که یک عمل چقدر خوب و چقدر بد است. این وظیفه ی عامل یادگیرنده است که با در دست داشتن این اطلاعات، یاد بگیرد که بهترین عمل در هر وضعیت کدام است. این موضوع، بخشی از نقاط قوت خاص یادگیری تقویتی است. از این طریق، مسائل پیچیده ی تصمیم گیری در اغلب اوقات می توانند با فراهم کردن کمترین میزان اطلاعات مورد نیاز برای حل مسأله، حل شوند. در بسیاری از حیوانات، یادگیری تقویتی، تنها شیوه ی یادگیری مورد استفاده است. همچنین یادگیری تقویتی، بخشی اساسی از رفتار انسان ها را تشکیل می دهد.

هنگامی که دست ما در مواجهه با حرارت می سوزد، ما به سرعت یاد می گیریم که این کار را بار دیگر تکرار نکنیم. لذت و درد مثالهای خوبی از پاداش ها هستند که الگوهای رفتاری ما و بسیاری از حیوانات را تشکیل می دهند. در یادگیری تقویتی، هدف اصلی از یادگیری، انجام دادن کاری و یا رسیدن به هدفی است، بدون آنکه عامل یادگیرنده، با اطلاعات مستقیم بیرونی تغذیه شود.



شکل ۱.۰: اصول یک سیستم یادگیری تقویتی

مرور سوابق پیشینه

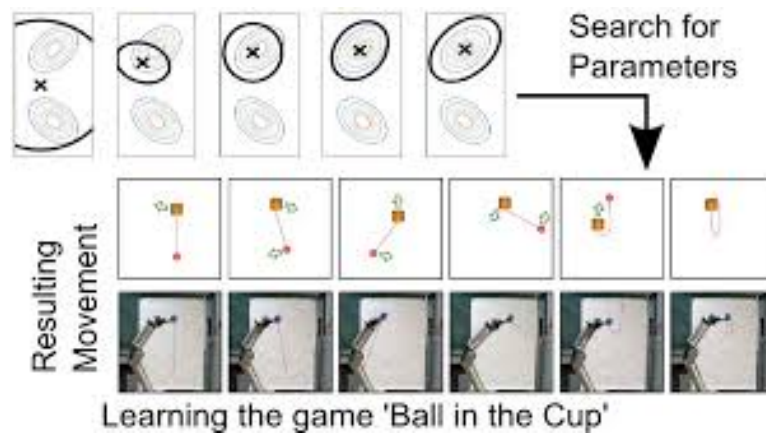
تحقیقات جاری:

- پیدا کردن راهکارهای قابل انطباق با تعداد کمتر (یا هیچ) پارامتری تحت شرط های بسیار زیاد.
- تخمین های تجربی بزرگ
- یادگیری و تصمیم گیری تحت اطلاعات محدود.
- یادگیری تقویتی سلسله مراتبی
- بهینه سازی راهکارهای «تابع - مقدار» و «جستجوی راهبرد» حاضر
- الگوریتم هایی که برای مجموعه بزرگ و حتی پیوسته تصمیمات ممکن کار کنند
- یادگیری های برای تمام عمر
- برنامه ریزی بهینه بر پایه نمونه (بر اساس درخت جستجوی مونت کارلو)
- یادگیری توزیع شده یا چند ماشینی
- استفاده از یادگیری تقویتی در زندگی واقعی

طرح پیشنهادی:

بررسی چند مسئله معروف در این زمینه:

- بازی flappy bird
- برج هانوی
- ۳*۳ روبیک
- کنترل پاندول معکوس
- مساله ی توپ و فنجان

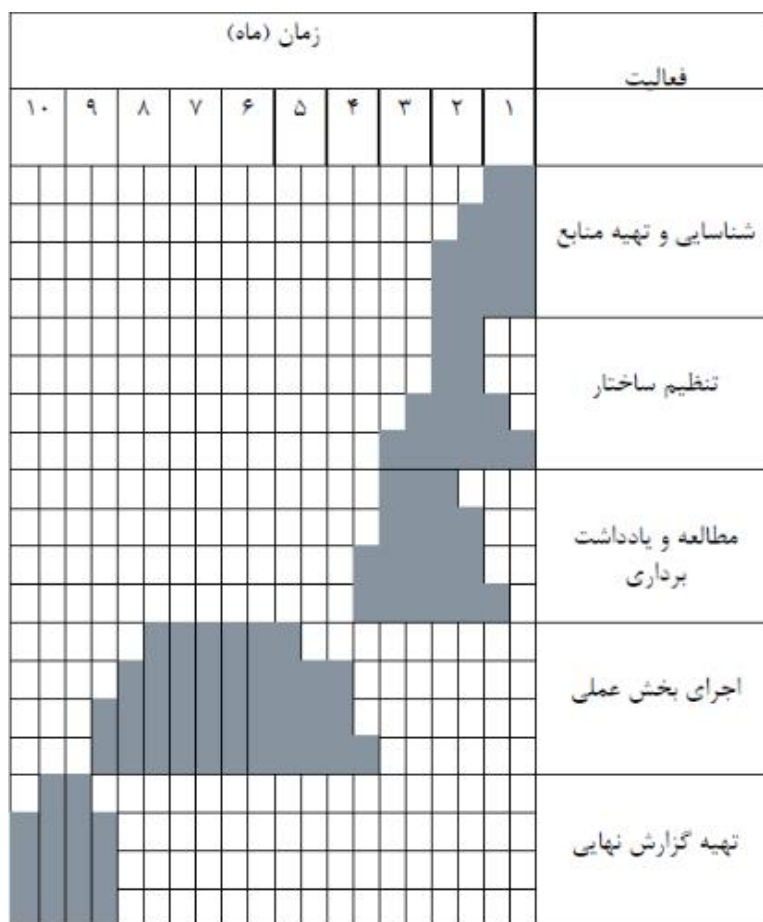


شکل ۲۰: مساله ی توپ و فنجان

مراحل انجام :

۱. شناسایی و تهیه ی منابع
۲. تنظیم ساختار
۳. تهیه ی بخشهای مقدمه، محتوای اصلی، نتیجه گیری و منابع
۴. مشخص کردن ترتیب مباحث محتوا
۵. تهیه ی فهرست مباحث اصلی و فرعی
۶. مطالعه و یادداشت برداری
۷. مطالعه ی مقالات اینترنتی
۸. مطالعه ی سایت های مفید در این باره
۹. مطالعه ی کتاب های مرتبط
۱۰. اجرای بخش عملی
۱۱. پیاده سازی یک الگوریتم ساده ی یادگیری ماشین با استفاده از زبان پایتون
۱۲. تهیه ی گزارش نهایی
۱۳. مستند سازی و نوشتن گزارش نهایی
۱۴. آمادگی جهت ارائه ی شفاهی

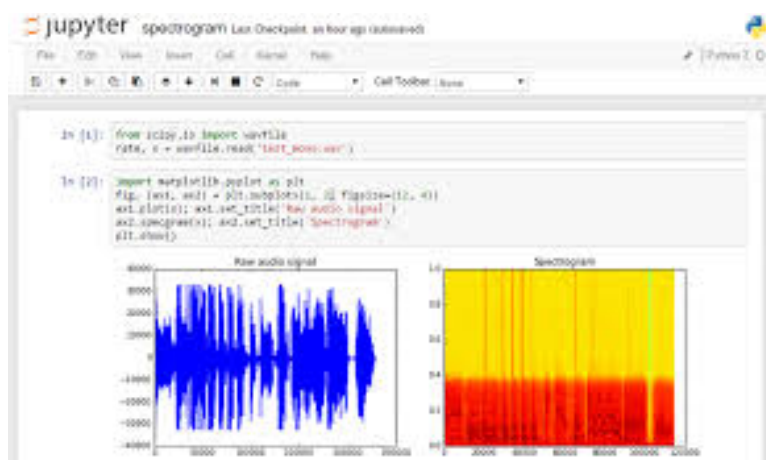
زمان بندی پروژه



شکل ۳.۰: زمان بندي پروژه

امکانات لازم

- کتابخانه های پایتون
- pycharm نرم افزار
- Jupyter notebook



شکل ۴.۰: jupyter notebook