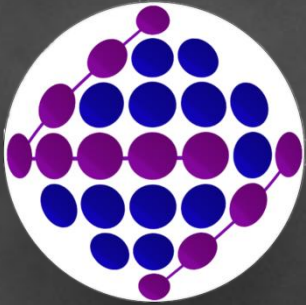


Reinforcement Learning



فصل نهم



شرکت فروش مصنوعي رسا

مدرس : فرید فاشمی نژاد

مشافعات OpenAI Gym



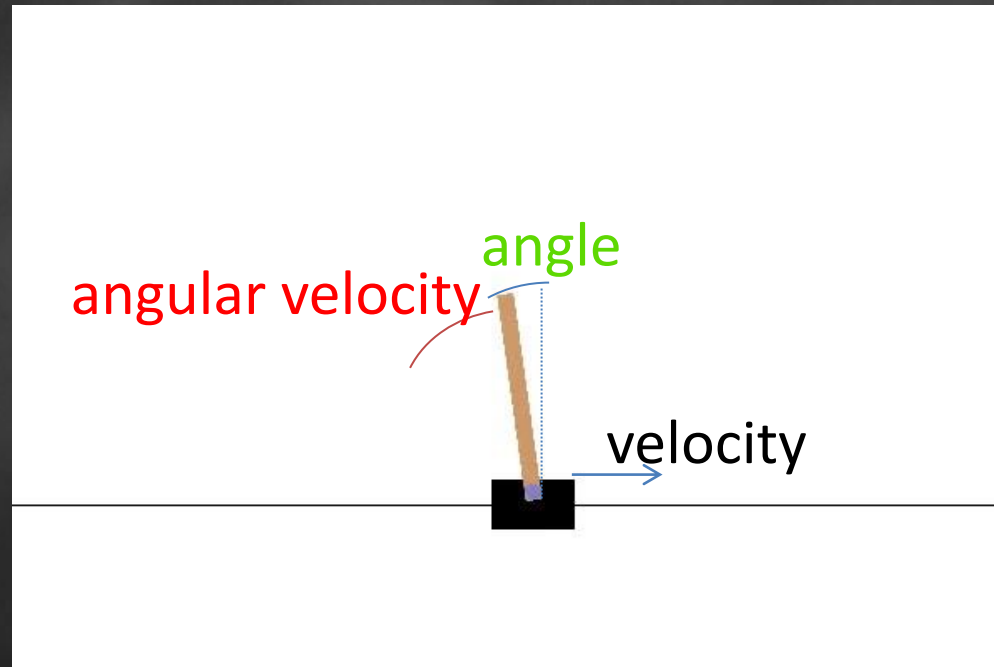
جلسه چهل و هشتم



شرکت فروش مصنوعی رسا

مدرس : خرید فاشمی نژاد

- تابع step اطلاعات مفید object ها را بر می گردانند



• Observations

• اطلاعات خاصی که از محیط باز میگردد،

به عنوان مشاهدات محیط یا

environment observations شناخته می

شود .

• مثال : زاویه، سرعت ، وضعیت بازی و . .

- جایزه یا reward

- مقدار جایزه به اقدام قبلی وابسته است.

- ابعاد تغییرات به محیط وابسته است اما agent همیشه میخواند به سطح پاداش بالاتر برسد

- Done

- متغیر بولینی نیاز به ریست شدن محیط را اعلام میکند.

- مثل باختن بازی، افتادن ستون و . . .

Info

-
-

دیکشنری که اطلاعات تشخیصی را
دارد و عموماً برای دیپاک استفاده می
شود.