# گزارش کد تمرین HW12 – محمد اصولیان 99521073

## چرا در حالت manual هدف به سمتی که شما حرکت میدهید نمیرود؟

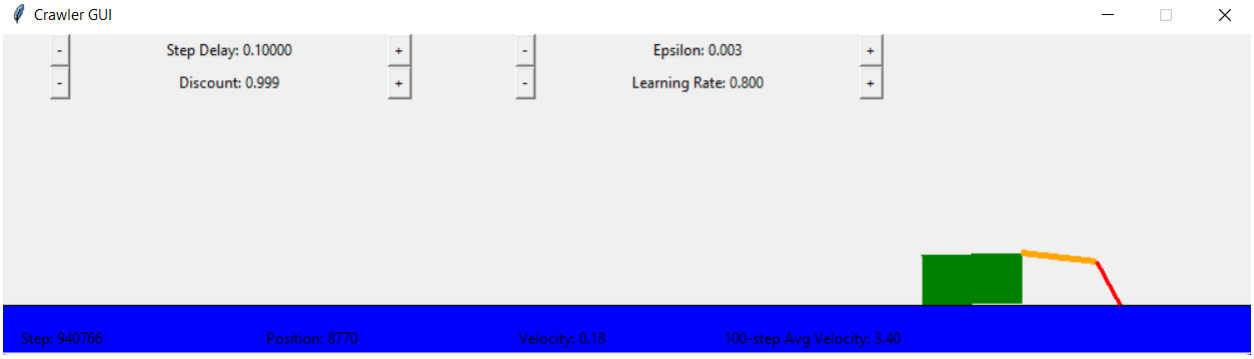
به علت وجود نویز در حرکت. در واقع ما فقط اکشن را تعیین میکنیم. این که طبق اکشن وارد شده چه transition ای انجام شود بستگی به احتمال T(s, a, s') دارد.

## تصویر لاگ و نتیجه نهایی و برداشت خود را ذکر کنید


نقطه آبی در ابتدا دانش اولیه ای ندارد بنابراین حرکات رندوم انجام میدهد. در مراحل اولیه که نقطه نمیداند حرکت optimal چیست و فقط رندوم عمل میکند، خیلی طول میکشد که به terminal state برسد. اما به مرور دانش ما از Qvalue ها بیشتر میشود و حرکات optimal به وجود می‌آیند. برای همین نقطه سریع تر به مقصد میرسد.

پس از چند اپیزود، حرکات optimal مشخص شده اند و در واقع مقادیر pi، converge شده اند. اما تا چند اپیزود بعد هم نقطه آبی روی همان مسیر سبز یافت شده حرکت میکند و مقادیر Q را آپدیت میکند. در واقع pi خیلی زود تر از Q، converge میشود.

## پارامتر های موجود در کد crawler را شرح دهید




step delay: زمان انجام هر step. با کم کردن این پارامتر میتوانیم مراحل را خیلی سریع تر جلو ببریم و learning را سرعت ببخشیم.

Discount: همان گاما در فرمول Q است و میزان اهمیت به value های نزدیک تر را بیان میکند. با کم کردن گاما خزنده حرکات کوتاه و سریع بر میدارد و جهش ها را ریز ریز بر میدارد. اما با زیاد کردن گاما بازوی خزنده در هر حرکت مسافت زیادی را برای خزنده طی میکند.

Learning Rate: همان آلفا در فرمول Q است. میزان اهمیت به سمپل های جدید را بیان میکند. با کم کردن آلفا سمپل های جدید تاثیر کمتری روی داده ها خواهند گذاشت.

Epsilon: مربوط به فرمول e-Greede است. هر چه میزان epsilon بیشتر باشد خزنده حرکات رندوم بیشتری انجام میدهد. و هر چه کم تر باشد، خزنده حرکات رندوم کم تری انجام میدهد و به جای آن حرکاتی را انجام میدهد که optimal باشند و value بیشتری برای خزنده به همراه داشته باشند.

در نوار پایین هم Step تعداد مراحل اجرا شده، Position جایگاه خزنده روی زمین، Velocity سرعت خزنده و 100-step Avg Velocity میانگین سرعت خزنده در 100 حرکت قبل را نشان میدهد.

میتوانیم با کم کردن step delay و زیاد کردن epsilon ابتدا ربات را به حالت learning ببریم و پس از مدتی epsilon را صفر کنیم تا ربات با تمام داده ای که به دست آورده در optimal ترین حالت خود کار کند.

My max 100-step Avg Velocity: 3.5