

گزارش کار پروژه سوم هوش

سوال ۱- در این بخش ما برای هر state تمام action ها را چک می کنیم تا تمام iteration های مجاز برای state ها را بررسی کنیم. در این بخش برای محاسبه action ها از q-value استفاده می کنیم تا بیشترین مقدار را برای state پیدا کنیم که به این منظور ما از دو تابع `computeQValueFromValue()` و `computeActionsFromValue()` و تکمیل آن استفاده می کنیم.

سوال ۲- در این سوال باید noise را برابر ۰ قرار دهیم تا به جواب برسیم.

سوال ۳- برای اینکه به خروجی نزدیک برسیم باید discount را افزایش داده و livingReward را منفی می دهیم همچنین برای خروجی های دور مقدار livingReward مثبت است همچنین برای اجتناب از صفر discount قرار می دهیم

سوال ۴- در این بخش با توجه به در هر value, iteration های اکشن های مربوط به state مورد نظر را بررسی می کنیم و value آنرا آپدیت می کنیم.

سوال ۵- در این بخش نیز با توجه به توضیحات تابع `runValueIteration` را کامل می کنیم. در این قسمت یک priority queue می سازیم و posterior های یک state را نگهداری می کنیم تا حالت تکراری به وجود نیاید. بر روی همه state های ممکن فرایند مارکوف را پیمایش می کنیم و state را به priority queue اضافه می کنیم.

سوال ۶- در این بخش می خواهیم Q-learning را پیاده سازی کنیم برای اینکار در ابتدا یک counter برای q-value تعریف می کنیم سپس تابع `computeValueFromQValues` با انجام پیمایش روی اکشن های state های ما بیشترین مقدار برگردانده شده را به counter ما برمی گرداند که آن به عنوان value ما است. همچنین در تابع `computeActionFromQValues` بهترین اکشن را از بین از بین q-value ها انتخاب می کنیم و در صورتی که چندین مقدار برابر داشته باشیم از بین آنها به صورت random انتخاب می کنیم.

سوال ۷- در این بخش با استفاده از تابع `flipcoin` اکشن های با احتمال epsilon و یا سیاست مناسب را به عنوان اکشن برمی گردانیم و این عمل را تا زمانی که همگرا شویم ادامه می دهیم.

سوال ۸- با اپسیلون برابر با ۱ اینکار نشدنی است پس باید NOT POSSIBLE رو return کنیم.

سوال ۹- با توجه به پیاده سازی که برای Q-learning کردیم در این بخش عامل پکمن با تجربه کردن state های مختلف رفعت رفته یاد می گیرد تا از روح ها دوری کند.

سوال ۱۰- در تابع `getQValue` weight را در feature ضرب می کنیم و در تابع `update` features را حساب می کنیم و diff را با توجه به معادله گفته شده حساب می کنیم و در نهایت مقدار weight را update می کنیم.