

سوال ۲ عملی:

بخش آ)

فایل qlearningAgents.py تکمیل شده است.

بخش ب)

اگر از PacmanQAgent برای نقشه های mediumGrid و mediumClassic استفاده کنیم چون فضای state های این نقشه ها خیلی بزرگ است و همچنین تعداد time step های هر episode زیاد بوده، به تعداد بیشتری training episode نیاز داریم تا مقادیر $Q(s, a)$ برای موارد مختلف state و action به مقدار خوبی update شوند. بنابراین عملاً استفاده از این روش tabular برای این نقشه ها به صرفه نبوده و منابع زیادی برای train کردن آن نیاز است، این در حالیست که بسیاری از state ها وضعیتی نزدیک به هم دارند و از این موضوع در روش tabular Q-learning استفاده نمی شود.

بخش ج)

رفتار عجیب Pacman این است که نسبت به خوردن نقطه های بزرگ که در ۴ گوشه نقشه قرار دارند بی تفاوت بوده و وقتی نیز آن ها را می خورد و حالت روح ها تغییر می کنند باز هم از آن ها دوری کرده و نسبت به حالت روح ها نیز بی تفاوت است. این درحالی است که در صورتی که با خوردن آن نقطه های بزرگ و سپس خوردن روح ها می تواند امتیاز خود را ماکسیمایز کند. با کمی دقت در Feature هایی که توسط SimpleExtractor گزارش می شوند علت این موضوع را در می یابیم. در این Feature ها برای حالت روح ها و همچنین خوردن آن نقطه های بزرگ مقداری در نظر گرفته نشده و گزارش نمی شوند بنابراین عامل ما نسبت به این موارد بی تفاوت است.