

گزارش تمرین سری ۱۱ هوش مصنوعی - محمدپویا تراشی - ۴۰۱۵۲۱۱۰۲

توضیحات اجرا: برای تجربه بهتر از انیمیشن رسیدن عامل به هدف و نقشه حرارتی، فایل با پسوند py. اجرا گردد. برای بررسی کد، فایل با پسوند ipynb. گزینه مناسبتری است.

در این تمرین دو کلاس اصلی MazeGrid و Agent وجود دارد. وظیفه هر کدام به ترتیب ساخت محیط بازی و آموزش عاملی است که باید بازی یاد بگیرد.

کلاس MazeGrid:

توابع اصلی موجود:

- generate_maze: وظیفه ساخت محیط بازی و قرار دادن اجزا مانند دیوار، عامل و هدف را انجام می‌دهد.
- get_next_state: با گرفتن حرکت انجام شده و انجام آن، خانه بعدی‌ای که عامل در آن قرار می‌گیرد را مشخص می‌کند.
- is_valid_move: بررسی می‌کند حرکتی که در تابع get_next_state انجام می‌گیرد معتبر است یا خیر. در صورتی که از محیط خارج شد یا با دیوار برخورد داشت مانع از انجام حرکت می‌گردد.
- get_reward: تابع پاداش.

کلاس Agent:

توابع اصلی موجود:

- choose_action: در حالتی که عامل در حال آموزش نیست، یک حرکت تصادفی از بین چهار جهت را بر می‌گرداند. در غیر این صورت حرکت با بیشترین q_value را برمی‌گرداند.
- train: به تعداد اپیزودها، طبق الگوریتم Sarsamax یک اکشن را انتخاب می‌کند و مقدار q_value را بروز می‌نماید. در هر اپیزود تا زمانی ادامه می‌دهد که به terminal state برسد یا از محدودیت تعداد عبور کند.
- get_best_path: در حالتی که اپیزودهای آموزش تمام می‌شود، بیشترین q_value را از نقطه شروع تا پایان باز می‌گرداند.

سایر توابع جانبی مثل animate_path و visualize_q_values وظیفه نمایش نهایی را بر عهده دارند.

نمودار پاداش تجمعی:

این بازی به وسیلهٔ عامل در اپیزودهای بین ۱۰ تا ۱۰۰۰ تایی انجام گردید و مجموع پاداش‌های تجمعی ثبت شد. نمودار این پاداش‌ها به شکل زیر می‌باشد:

