



## تمرین شماره ۱ - درس هوش مصنوعی - پاییز ۱۴۰۱

دانشکده ریاضی، آمار و علوم کامپیوتر - دانشگاه تهران

### ۱ - شرح مسئله

هدف از ارائه این تمرین، آشنایی و پیاده سازی مسئله ی حل بازی مارپیچ با استفاده از الگوریتم Q-Learning است. در این مسئله محیطی به برنامه داده میشود. محیط شامل یک سری راه های قابل پیمایش است که با رنگ سفید مشخص شده است. برای گسسته سازی و دیجیتالی کردن راه ها، آنها را به مربع های واحد کوچک تبدیل میکنیم. محل شروع حرکت agent نیز مشخص میشود. هدف خارج شدن از این محیط است که با رسیدن به نقطه پایانی محقق میشود. همچنین مهم است که عامل از تمامی پرچم های موجود در محیط گذر کند. در شکل زیر موانع با رنگ قرمز مشخص شده اند. که W بیانگر Way یا راه، B بیانگر Block یا دیوار، F به معنی Flag یا پرچم، A به معنی Agent یا عامل و T به معنی Target یا هدف است.

A	B	W	W	W	W	W	W	W	W
W	W	W	F	W	B	W	W	W	W
F	W	W	W	W	B	F	W	W	W
B	B	W	B	B	W	B	W	W	W
W	F	B	W	B	F	B	B	B	W
W	W	B	W	B	W	W	W	W	W
W	W	W	W	W	W	W	W	W	W
W	F	W	W	W	W	B	B	B	B
W	B	B	B	B	B	W	W	W	W
W	W	W	W	W	W	W	B	F	T

- توضیح دهید تعداد حالت ها (State) وابسته به چه چیزی است؟ آیا میتوان تعداد حالت ها را کاهش داد؟
- State، Action، Rewards و Goal State را برای مسئله تعریف کنید و توضیحات خود را ارائه کنید. تعاریف شما باید به گونه ای باشد که پیمایش های شما منجر به پیمایش تمامی پرچم ها شود و روی Block ها نرود. همچنین در پایان بر روی T یا همان Target قرار بگیرد.
- در ابتدا محیط را به صورت شکل بالا در نظر بگیرید. به ازای حداقل ۳ مقدار  $\gamma$  برابر ۰/۲۵، ۰/۵ و ۱ تاثیر  $\gamma$  را تحلیل کنید. همچنین به ازای حداقل ۳ مقدار  $\alpha$ ، نتایج را تحلیل کنید و تاثیر  $\alpha$  را مشخص کنید.
- محیط یا environment را به صورت یک گراف وزن دار نمایش دهید. رئوس یال حالت های مختلف و وزن های آن میزان Reward بعد از اعمال Action مربوطه است. می توانید روی هر یال Action مربوطه را به عنوان یک Label ذکر کنید.

- Q-Table را برای این مسئله را رسم کنید.
- Episode بدست آمده را به همراه تغییرات reward برای این مسئله را رسم کنید.
- کد های خود را روی یک محیط دلخواه دیگر آزمایش کنید و نتیجه را بیان کنید.
- (افزونه/امتیازی)\* نمایش episode ها و reward های حاصل از اعمال مختلف در هر حالت به صورت گرافیکی (تصویر یا انیمیشن) امتیاز مثبت و اختیاری خواهد بود. برای نمایش گرافیکی میتوانید از پکیج های آماده موجود برای پایتون مثل PyGame، Tkinter، PyQt5 و ... استفاده کنید.
- (افزونه/امتیازی)\* فرض کنید در بعضی از مربع های واحد، مانع متحرکی قرار دارد. این موانع با O بیانگر Obstacle مشخص شده‌اند. Agent می‌تواند در صورتی که سمت دیگر مانع خالی (همان W) باشد، مانع را حل داده و از سر راه کنار بزند (توجه کنیم که مانع در جهت حرکت Agent حرکت می‌کند). برای مثال سناریو زیر را در نظر بگیرید:

W	B	W	W	W	W	W	W	W	W
W	W	W	F	W	B	W	W	W	A
F	W	W	W	W	B	F	W	W	O
B	B	W	B	B	W	B	W	W	W
W	F	B	W	B	F	B	B	B	W
W	W	B	W	B	W	W	W	W	W
W	W	W	W	W	W	W	W	W	B
W	F	W	W	W	W	B	B	B	B
W	B	B	B	B	B	W	W	W	W
W	W	W	W	W	W	W	B	F	T

W	B	W	W	W	W	W	W	W	W
W	W	W	F	W	B	W	W	W	W
F	W	W	W	W	B	F	W	W	A
B	B	W	B	B	W	B	W	W	O
W	F	B	W	B	F	B	B	B	W
W	W	B	W	B	W	W	W	W	W
W	W	W	W	W	W	W	W	W	B
W	F	W	W	W	W	B	B	B	B
W	B	B	B	B	B	W	W	W	W
W	W	W	W	W	W	W	B	F	T

- می‌توان مشاهده کرد که agent مانع را یک واحد به سمت پایین حل داده است. اما توجه کنیم که با این کار باعث شده دیگر مسیری برای رسیدن به هدف وجود نداشته باشد. در نتیجه agent باید از چنین حرکاتی

بپرهیزد. همچنین ممکن است بدون حرکت دادن موانع متحرک هیچ مسیری برای رسیدن به هدف وجود نداشته باشد. آیا می‌توانید برای این مسئله راه حل پیشنهاد دهید؟

## ۲ - نحوه ارسال

- فایل توضیحات خود (به صورت pdf) را به همراه فایل کد ها که ترجیحا در notebook با پسوند .ipynb است را در پوشه ای قرار داده و پوشه را zip کنید.
- نحوه نامگذاری فایل شما باید به صورت Assignment1 – YourName – AI1401 باشد. پاسخ ها را در سایت elearn آپلود کنید.
- استفاده از همه ی زبان های برنامه نویسی بلامانع است اما ترجیحا از Python استفاده کنید.