

# به نام خدا

پروژه هوش مصنوعی

مهرداد قصابی

۹۷۳۶۱۳۰۶۰



# پیمان اخلاقی

بنده تعهد می نمایم که پروژه تحویل داده شده نتیجه کار خویش بوده  
و در هیچ یک از بخش های پروژه از کار کسی کپی برداری نشده است  
در صورتی که مشخص گردد این پروژه کار بنده نبوده است طبق ضوابط آموزشی  
با من برخورد شده و حق اعتراضی نخواهم داشت

# مسیریابی

زمان مناسب یا مناسب ترین مسیر ؟

چالش اصلی ما برای انتخاب الگوریتم مناسب مسیریابی این است که بایستی بین پیدا کردن بهترین جواب یا یک جواب با زمان اجرای مناسب انتخاب کنیم.

با در نظر گرفتن صفحه بازی به عنوان گراف میتوانیم. اگر که خانه های جدول بازی را به عنوان نود های گراف در نظر بگیریم با توجه به آن که عامل بازی در هر خانه فقط حق حرکت به همسایه های وان نیومن خود را دارد پس هر نود گراف ما دارای چهار همسایه است

سه شیوه مسیریابی در ادامه مطرح میشوند که ما نسبت به زمان خود بین آن ها تغییر راهبرد میدهیم.

# مسیر یابی بهترین اول

در الگوریتم حریصانه بهترین اول ایده کلیدی این است که ما یک حلقه در حال گسترش به نام مجموعه پیشرو (فرانتیر) داریم که این مجموعه بایستی به طور برابر در تمامی جهات گسترش پیدا کند.

این الگوریتم تنها در مسیر های امیدوار کننده برای جواب کاوش میکند بنابراین شاید بهترین جواب پیدا نشود.

راهبرد ما اینگونه است که از این الگوریتم وقتی استفاده کنیم که زمان اجرای مناسبی در اختیار نداریم و پیدا کردن یک جواب مناسب در زمان کوتاه را به بهترین جواب ترجیح میدهیم

# الگوریتم دایجکسترا

در این الگوریتم ما بایستی اولویت بندی کنیم که کدام مسیرها را کاوش کنیم. این الگوریتم به جای کاوش در همه مسیرهای ممکن به طور مساوی، مسیرهای کم هزینه را ترجیح می دهد. با تعریف مناسب تابع هزینه ما میتوانیم مسیر درست را پیدا کرده و از ایجاد جنگل جلوگیری کنیم

وقتی ما زمان اجرای مناسبی در اختیار داریم ما میتوانیم از این الگوریتم به جای الگوریتم حریصانه عرض اول استفاده کنیم تا با یافتن بهترین مسیر از زمان خود استفاده کرده باشیم

# الگوریتم ای استار

الگوریتم دایجکسترا برای یافتن کوتاه‌ترین مسیر به خوبی کار می‌کند، اما زمان را برای کاوش در جهت‌هایی که امیدوارکننده نیستند تلف می‌کند.

جستجوی حریصانه عرض اول نیز که در مسیرهای امیدوارکننده کاوش می‌کند اما ممکن است کوتاه‌ترین مسیر را پیدا نکند.

همانطور که گفته شد راهبرد الگوریتم ما این است که وقتی ما زمان اجرای مناسبی در اختیار داریم از الگوریتم دایجکسترا و وقتی زمان مناسبی نداریم از الگوریتم بهترین اول استفاده می‌کنیم در این میان وجود یک الگوریتم میانجی برای حالت‌های متوسط ضروری به نظر می‌رسد

بنابراین ما در این پروژه از الگوریتم ای استار استفاده می‌کنیم که ویژگی‌های مطلوب هر دو الگوریتم بالا را تا حدودی داشته باشد

# تابع اکتشافی

در الگوریتم های یاد شده مجموعه پیش رو در تمامی جهات گسترش می یابد  
اما آیا این کار بهینه است؟

مطمئناً اگر کاری انجام دهیم که مجموعه پیشرو به سمت درستی پیش رود در خیلی از هزینه ها صرفه جویی  
میگردد.

در صورتی که تابع اکتشافی را فاصله منتهن نود کنونی و نود مقصد تعریف کنیم تقریب خوبی برای یافتن  
مقدار تقریبی هزینه داریم

بنابراین یک نود در گراف بایستی همسایه وان نیومنی را برای حرکت انتخاب کند که جمع هزینه انتقال به آن و  
تابع اکتشافی رسیدن به مقصد کمتر است

# تابع انتخاب

نوشتن یک تابع خوب برای اولویت بندی کردن هدف ها از هر چیزی ضروری تر به نظر میرسد.

میتوان این تابع را با نسبتی از فاصله و امتیاز یک هدف نوشت که توان های صورت مخرج نمایانگر ترجیح امتیاز و فاصله به یکدیگرند

اما شاید نگاشتن یک تابع ثابت که عامل همیشه به آن استناد کند چندان بهینه به نظر نمیرسد بنابراین عامل بایستی یاد بگیرد که به چه هدف هایی گرایش پیدا کند (در فاز یک کیو لرنینگ کامل پیاده سازی نشده)



# کیو لرنینگ

کیو لرنینگ یک شیوه تقویتی یادگیری ماشین است که از یک  
ارایه دو بعدی به روزرسانی شونده برای یادگیری استفاده  
میکند که هر خانه در این ارایه نماینگر امتیاز انتخاب یک  
کنش در یک مرحله خاص است

$$Q^{\pi}(s_t, a_t) = \underline{E}[\underline{R}_{t+1} + \gamma \underline{R}_{t+2} + \gamma^2 \underline{R}_{t+3} + \dots | s_t, a_t]$$

ارزش Q برای حالت یا توجه به عمل

پاداش تجمعی تخفیف داده شده مورد انتظار

حالت و عمل داده شده

## منابع

<https://www.redblobgames.com/pathfinding/a-star/introduction.html>

<https://blog.faradars.org/reinforcement-learning-and-q-learning/>