به‌نام خدا

هوش مصنوعی

تمرین ۱ – تئوری

مهراد میلانلو | ۹۹۱۰۵۷۷۵

استاد:

دکتر رهبان

سوال ۱.

آ)

درست؛ برای مثال در یک محیط تک‌حالته که تمام اکشن‌ها cost و reward یکسان دارند، انتخاب میان رفتارها هیچ تفاوتی ندارد و بنابراین تمام عوامل عاقلانه رفتار می‌کنند. (اثبات وجودی است)

ب)

درست؛ در یک محیط، برای یک ایجنت که عقلانی رفتار می‌کند، تعدادی بخش غیرقابل دسترسی وجود دارد. بنابراین با تغییر دلخواه این بخش‌ها به‌طور دلخواه، تغییری در رفتار ایجنت به‌وجود نمی‌آید و آن بخش‌ها همچنان Unreachable هستند و بنابراین در دو محیط مختلف ایجنت عقلانی رفتار می‌کند.

ج)

نادرست؛ تصمیم‌گیری کاملا عقلانی بدان معناست که ایجنت با توجه به اطلاعات دریافتی‌اش بتواند تصمیم‌گیری خوب و مثبتی داشته باشد.

سوال ۲.

آ)

اندازه‌ی فضای مسئله در واقع تعداد تمام استیت‌های ممکن در مسئله است. از آن‌جایی که MN خانه در صفحه وجود دارد، می‌توانیم با درنظر نگرفتن بلوک سیاه برای به‌دست آوردن کران بالا (چون ممکن است از این خانه‌ها مطلع نباشیم تعدادشان را ۰ در نظر می‌گیریم) بگوییم برای قرارگیری ماشین‌ها در صفحه C(2,MN) حالت داریم. برای خانه‌های دیگر نیز یا آتش در آن‌ها هست و یا نیست. چون در

خانه‌هایی که ماشین هست نمی‌توانیم آتش داشته باشیم، بنابراین در نهایت کران بالای اندازه‌ی مسئله برابر است با: (دقت کنید تخمین‌های بهتری نیز می‌توان داد)

$$|states| = C(2,\ MN) \times 2^{(MN-2)}$$

ب)

چون در هر مرحله هر ماشین می‌تواند در چهار جهت حرکت کند یا اصلا حرکت نکند، طبق اصل ضرب، برای کران بالای ضریب انشعاب خواهیم داشت:

$$b \leq 5 \times 5 = 25$$

ج)

می‌دانیم توابع اکتشافی یک‌نوا زیرمجموعه‌ی سره مجموعه توابع قابل‌قبول هستند. پس اگر تابع اکتشافی انتخابی، یک‌نوا باشد، قابل‌قبول نیز هست. به‌همین دلیل تابع هیوریستیک خود را «فاصله‌ی منهتن ماشین‌ها با نزدیک‌ترین هدف» در نظر می‌گیریم. مقدار این تابع در حالت هدف صفر است. در سایر حالات نیز چون حرکت مورب نداریم، در هر حالت حداکثر یک واحد فاصله کاهش می‌یابد و از سوی دیگر هزینه‌ی حرکت بین حالات نیز یک واحد است. بنابر نامساوی مثلثی این تابع اکتشافی یک‌نوا است.
به بیان ساده‌تر می‌توان گفت طبق تعریف Admissable بودن تابع Heuristic، چون حرکت مورب نداریم پس در هر مرحله:

$$H(n) \leq H^{*}(n)$$

پس تابع قابل‌قبول است. از سوی دیگر چون برای هر استیت، تابع اکتشافی عدد ثابت نیست و به فاصله‌ی ایجنت از هدف بستگی دارد، بنابراین تابع غیربدیهی‌است.
دقت کنید توابع اکتشافی بهتری هم می‌توان ارائه داد. برای مثال یک‌بار ماکسیمم فاصله‌ی منهتن A و B با هدف ۱ و ۲ را به ترتیب محاسبه کرده و بار دیگر ماکسیمم فاصله‌ی منهتن A و B با هدف ۲ و ۱ را به ترتیب محاسبه کنیم. مینیمم این دو مقدار می‌تواند تابع اکتشافی باشد که با همان استدلال بالا قابل قبول است.

سوال ۳.

آ)

می‌دانیم تابع f محدب است اگر و فقط اگر مشتق دوم آن ("f) مثبت باشد. بنابراین مشتق اول آن ('f) همواره صعودی‌است. چون تابع اکتشافی قابل قبول است، داریم:

$$h(x) \ \leq \ h^{*}(x) \ \Rightarrow \ f'(h(x)) \ \leq \ f'(h^{*}(x))$$

بنابراین حکم اثبات می‌شود.

ب)

به‌ازای هرنقطه، تابع $h_x(n)$ را برابر ماکسیمم تمام توابع دیگر درنظر می‌گیریم. چون تمامی توابع قابل‌قبول هستند، پس $h_x(n)$ هم قابل قبول است و چون در هر نقطه بیشتر از بقیه‌ی نقاط است، تمام توابع دیگر را غالب می‌کند.

$if\ h_1(x) \ \leq \ h_2(x)\ for\ all\ x\ (both\ are\ admissable) \ \Rightarrow \ h_2(x)\ Dominates\ h_1(x)$

$h_x(n) \ = \ \forall m \in R:\ h_x(m) \ = \ max(h_1(m),\ h_2(m),\ ...,\ h_m(m)) \ \leq h^{*}(m)$

$\Rightarrow h_x(n)\ is\ admissable.$

$\Rightarrow h_x(n) \ \geq h_1(n),\ h_2(n),\ ...,\ h_m(n) \ \Rightarrow \ h_x(n)\ Dominates\ h_1(n),\ h_2(n),\ ...,\ h_m(n)$

سوال ۴.

روش Genetic شیوه‌ی پیاده‌سازی یکسانی در مکانیزم‌ها ندارد و می‌توان از شیوه‌های مختلف بهره برد. برای مثال می‌توان این‌گونه عمل کرد:

مسئله‌ی SSP انتخاب زیرمجموعه‌ای از یک مجموعه است به‌طوری که مجموع اعداد زیرمجموعه‌ی انتخاب شده تا حد امکان نزدیک به مقدار معین T (یا حتی برابر) باشد. بنابراین برای استفاده از الگوریتم ژنتیک فرض می‌کنیم طول مجموعه‌ی داده شده متناهی‌است و یک رشته‌ی باینری مانند S

به‌طول اندازه‌ی مجموعه‌ی اصلی در نظر می‌گیریم. هر بیت در S را برابر ۱ قرار می‌دهیم اگر عدد متناظر با آن در زیرمجموعه‌ی انتخابی باشد و در غیر این‌صورت ۰ قرار می‌دهیم.

مکانیزم Crossover را بعد از انتخاب دو رشته (Selection) انجام می‌دهیم. سپس برای هر رشته یک عدد تصادفی کوچکتر از طول رشته انتخاب کرده و از آن‌جا یک زیررشته به‌طول دلخواه (برای مثال هر کدام نصف طول رشته‌ی اصلی) انتخاب می‌کنیم. سپس زیررشته‌ی یک‌کدام را جایگزین باقی‌مانده‌ی رشته‌ی دیگر می‌کنیم. متناظرا می‌توان این کار را در رشته‌ی دیگر نیز انجام داد و دو رشته‌ی جدید در جامعه‌ی رشته‌ها خواهیم داشت.

مکانیزم Mutation یا جهش را پس از Crossover انجام می‌دهیم. در این‌جا می‌توانیم به‌طور تصادفی یک یا چند بیت از رشته را انتخاب کرده و مقدار آن را Toggle کنیم. این عمل را برای اضافه کردن پارامتر تصادفی به الگوریتم و از دست ندادن شانس رسیدن به جواب درست بعد از Crossoverهای اشتباه، انجام می‌دهیم.

سوال ۵.

آ)

در حالت انتخاب چرخ رولت، احتمال انتخاب هر کروموزم f_i برابر است با:

$$Pr[f_i] = \frac{f_i}{\sum_j^n f_j}$$

$$i = 4 \rightarrow Pr[f_4] = \frac{10}{5+7+8+10+15} = \frac{10}{45} = \frac{2}{9} = 0.22$$

ب)

در این حالت به بزرگترین f_i امتیاز ۱۰ و به کوچکترین f_i امتیاز ۱ می‌دهیم. سپس امتیاز باقی فیتنس‌ها را به‌صورت خطی مشخص می‌کنیم. (در واقع دنبال معادله‌ی خطی هستیم که f_1 را به f_1' و f_5 را به f_5' مپ کند. بنابراین خواهیم داشت:

$f(x) = ax + b$

$f(f_1) = af_1 + b \Rightarrow 5a + b = 1$

$f(f_5) = af_5 + b \Rightarrow 15a + b = 10$

$\Rightarrow a = \frac{9}{10}, b = \frac{-7}{2} \Rightarrow f(x) = \frac{9}{10}x - \frac{7}{2}$

$\rightarrow f'_1 = 1, f'_2 = 2.8, f'_3 = 3.7, f'_4 = 5.5, f'_5 = 10$

$i = 4 \rightarrow Pr[f_4] = \frac{5.5}{1+2.8+3.7+5.5+10} = \frac{5.5}{23}$

ج)

برای انتخاب یک تورنمنت با سایز ۲ از ۵ کروموزم، $C(2, 5) = 10$ حالت داریم. که سه حالت است:

۱- کروموزم ۴ در تورنمنت نباشد: $Pr_1(f_4) = 6 \times 0 = 0$

۲- کروموزم ۴ به‌عنوان کروموزم با فیتنس کوچک‌تر در تورنمنت باشد:

$Pr_2(f_4) = 1 \times 0.25 = 0.25$

۳- کروموزم ۴ به‌عنوان کروموزم با فیتنس بزرگ‌تر در تورنومنت باشد:

$Pr_2(f_4) = 3 \times 0.75 = 2.25$

با توجه به اصل جمع داریم:

$$Pr[f_4] = \frac{Pr_1(f_4)+Pr_2(f_4)+Pr_3(f_4)}{10} = \frac{0.25+2.25}{10} = 0.25$$

سوال ۶.

آ)

برای اثبات محدب بودن $F(x)$ کافی‌ست ثابت کنیم: $F''(x) \geq 0$.

$$F(x) = \frac{1}{x}\int_0^x f(t)dt, \quad x \in R_{++}$$

$$g(x) = \int f(x)dx$$

$$F(x) = \frac{1}{x}\int_0^x f(t)dt = \frac{1}{x}(g(x) - g(0))$$

$$F'(x) = \frac{d}{dx}F(x) = \frac{d}{dx}\left(\frac{g(x)-g(0)}{x}\right) = \frac{xf(x)-(g(x)-g(0))}{x^2} = \frac{f(x)}{x} + \frac{g(0)-g(x)}{x^2}$$

$$F''(x) = \frac{d}{dx}F'(x) = \frac{d}{dx}\left(\frac{f(x)}{x} + \frac{g(0)-g(x)}{x^2}\right) = \frac{xf'(x)-f(x)}{x^2} + \frac{-x^2f(x)-2x(g(0)-g(x))}{x^4}$$

$$= \frac{x^3f'(x)-2x^2f(x)-2x(g(0)-g(x))}{x^4} = \frac{x^2f'(x)-2xf(x)+2(g(x)-g(0))}{x^3} \quad (1)$$

$$\forall x \in R_{++}: x^3 > 0$$

$$h(x) = x^2f'(x) - 2xf(x) + 2(g(x) - g(0))$$

$$h(0) = 0 \quad (*)$$

$$h'(x) = 2xf'(x) + x^2f''(x) - 2f(x) - 2xf'(x) + 2f(x) = x^2f''(x)$$

$$f(x)\ is\ Convex \rightarrow f''(x) \geq 0 \rightarrow h'(x) \geq 0 \quad (**)$$

$$(*),\ (**) \rightarrow h(x) \geq 0;\ \forall x \in R_{++}$$

$$(1) \rightarrow F''(x) \geq 0;\ \forall x \in R_{++} \rightarrow F(x)\ is\ Convex.$$

ب)

طبق تعریف Convexity می‌دانیم:

$$\forall x, y \in R^{2}_{+}, \lambda \in [0, 1] : f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda)f(y)$$

$$t = \lambda b + (1-\lambda)a\,;\; t \in [a, b]\,;\; \lambda \in [0, 1]$$

$$\Rightarrow \frac{dt}{d\lambda} = b - a \Rightarrow dt = (b-a)d\lambda$$

$$\int_a^b f(t)dt = \int_0^1 f(\lambda b + (1-\lambda)a)(b-a)d\lambda = (b-a)\int_0^1 f(\lambda b + (1-\lambda)a)d\lambda$$

$$\le (b-a)\int_0^1 \lambda f(b) + (1-\lambda)f(a)d\lambda$$

$$= (b-a)[(\int_0^1 \lambda f(b)d\lambda) + (\int_0^1 (1-\lambda)f(a)d\lambda)]$$

$$= (b-a)[\frac{\lambda^2}{2}f(b)|^1_0 + \frac{2\lambda-\lambda^2}{2}f(a)|^1_0$$

$$= \frac{1}{2}(b-a)(f(b) + f(a))$$

$$\Rightarrow \int_a^b f(t)dt \le \frac{1}{2}(b-a)(f(b) + f(a)) \Rightarrow Proof\ is\ Completed.$$