هوش مصنوعی

نیم‌سال اول ۰۱-۰۰

مدرس: دکتر محمدحسین رهبان

دانشکده مهندسی کامپیوتر

تمرین هفتم ـ بخش اول

محمدجواد هزاره

شماره دانشجویی: ۹۸۱۰۱۰۷۴

سوال ۱

برای مدل‌سازی سه حالت زیر را در نظر می‌گیریم. با توجه به تقارن، هر حالت دیگری را می‌توان به شکل یکی از این سه حالت در نظر گرفت. G نشان دهنده‌ی مکان روح و P مکان پک‌من را نشان می‌دهد.

G	
	P

(s_1)

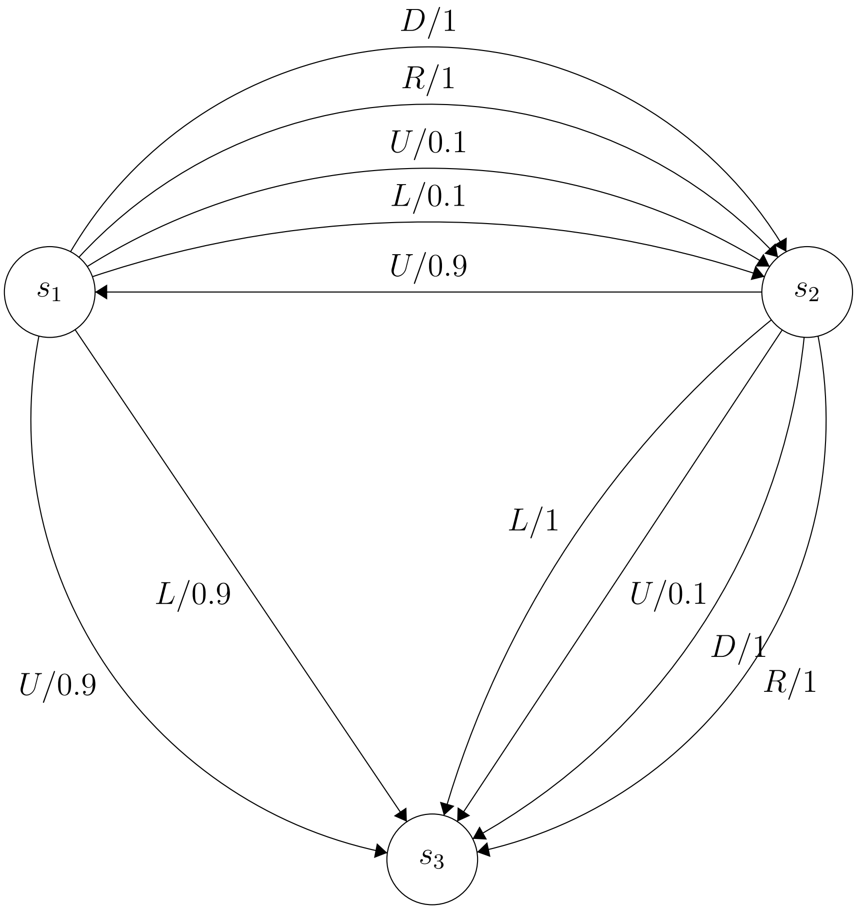
G	P

(s_2)

	G P

(s_3)

اگر فرض کنیم روح همواره به نحوی حرکت خواهد کرد که فاصله‌اش را با پک‌من کم کند، آن‌گاه برای مدل‌سازی مسئله خواهیم داشت: (برای هر یال، حرکت پک‌من و بلافاصله بهترین حرکت روح در نظر گرفته شده است.) reward مربوط به یال‌های منتهی به s_1 و s_2 برابر با 1 و سایر یال‌ها 0 خواهد بود.



آ) اگر داشته باشیم $x \in \{D, R\}$ ، $y \in \{U, L\}$ ، $z \in \{U\}$ و $w \in \{D, L, R\}$ ، با توجه به گراف مسئله، ارزش سیاست ما نه به خود اکشن‌ها بلکه به دسته‌ای که اکشن را از آن انتخاب می‌کنیم بستگی خواهد داشت. حال سیاست زیر را در نظر می‌گیریم:

$$\pi^* = \begin{cases} s_1 \to x \\ s_2 \to z \end{cases}$$

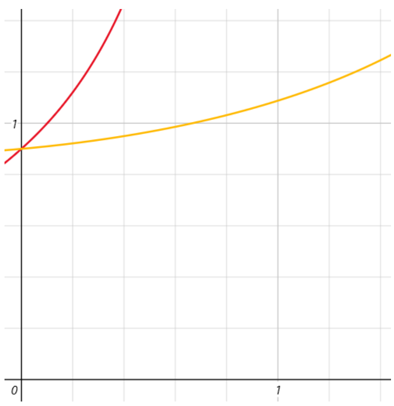
حالت s_3 نیز کنشی نخواهد داشت. ادعا می‌کنیم این سیاست خواسته‌ی مسئله را برآورده می‌کند. به عبارتی به ازای هر سیاست π، خواهیم داشت $\pi^* \geq \pi$. با توجه به تعریفی که برای x، y، w و z داشتیم، سه حالت دیگر برای π باقی می‌ماند که به صورت زیر هستند:

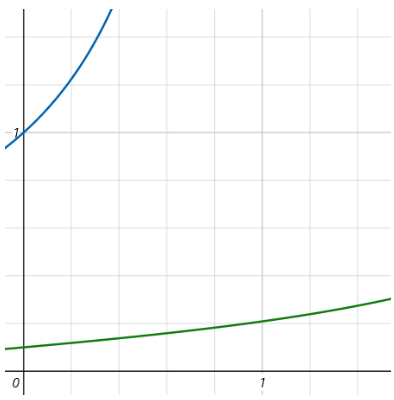
$$\pi_1 = \begin{cases} s_1 \to x \\ s_2 \to w \end{cases} \qquad \pi_2 = \begin{cases} s_1 \to y \\ s_2 \to z \end{cases} \qquad \pi_3 = \begin{cases} s_1 \to y \\ s_2 \to w \end{cases}$$

با حل معادلات تابع ارزش برای هر یک از این سیاست‌ها خواهیم داشت:

$$V^{\pi^*} = \begin{bmatrix} \frac{1+0.9\gamma}{1-0.9\gamma^2} \\ 0.9\frac{1+\gamma}{1-0.9\gamma^2} \end{bmatrix} \qquad V^{\pi_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad V^{\pi_2} = \begin{bmatrix} 0.1\frac{1+0.9\gamma}{1-0.09\gamma^2} \\ 0.9\frac{1+0.1\gamma}{1-0.09\gamma^2} \end{bmatrix} \qquad V^{\pi_3} = \begin{bmatrix} 0.1 \\ 0 \end{bmatrix}$$

که اگر γ عددی مثبت و کوچک‌تر از 1 باشد، مشخص است که π^* از π_1 و π_3 بهتر است. همچنین از π_2 بهتر خواهد بود که برای اثبات آن می‌توان به نمودارهای زیر دقت کرد: ($0.9\frac{1+\gamma}{1-0.9\gamma^2}$ ،$0.9\frac{1+0.1\gamma}{1-0.09\gamma^2}$ ،$\frac{1+0.9\gamma}{1-0.9\gamma^2}$ ،$0.1\frac{1+0.9\gamma}{1-0.09\gamma^2}$)





ب) با توجه به قسمت قبل، سیاست زیر خواسته‌ی مسئله را برآورده می‌کند:

$$\pi^*_{\text{det}} = \begin{cases} s_1 \to D \\ s_2 \to U \end{cases}$$