

به نام خدا

محمد مهدی آقا جانی

تمرین دوم

هوش مصنوعی

استاد نیک آبادی

پاییز ۹۵

تمرین ۱ :

معیار کارایی	محیط	عملگر ها	حسگر ها
کنترل چراغ راهنمایی رانندگی	کمینه کردن میانگین زمان انتظار همزمان با کمینه شدن تصادفات و کمینه کردن ترافیک	چهار راه و خیابان	چراغ های راهنمایی دوربین کنترل کننده ، حسگر های کف خیابان
تشخیص اسپم	درصد تشخیص صحیح اسپم	ایمیل های ارسال شده	توابع مربوط به تشخیص ایمیل
کنترل کننده ماشین لباسشویی	سریع ، تمیز ، کمترین آسیب به لباس	لباس ها	بازو های عملیاتی ماشین لباس شویی مانند پخش کننده مایع شوینده یا گرداننده اتاق ماشین
پیشنهاد دهنده دوست	میزان پایداری دوستی های پیشنهاد شده	شبکه های اجتماعی	توابع پیشنهاد دهنده دوست توابع کنترل کننده فعالیت ها و اشتراکات

قابل مشاهده	تک عامله / چند عامله	قطعی / تصادفی	مرحله ای / ترتیبی	ایستا / پویا	گسسته / پیوسته	شناخته / ناشناخته
کنترل چراغ راهنمایی	قابل کاملاً مشاهده	چند عامله	قطعی	ترتیبی	پویا	پیوسته شناخته
تشخیص اسپم	قابل کاملاً مشاهده	تک عامله	قطعی	مرحله ای	ایستا	گسسته شناخته
کنترل کننده ماشین لباس شویی	قابل کاملاً مشاهده	تک عامله	قطعی	مرحله ای	نیمه پویا	پیوسته شناخته
پیشنهاد دهنده دوست	جزیی قابل مشاهده	تک عامله	تصادفی	ترتیبی	پویا	پیوسته شناخته

تمرین دوم :

الف (بله . چون برنامه عامل شامل الگوریتم و داده میباشد و الگوریتم نیز میتواند متفاوت باشد مثلا تابع عامل برای مساله مرتب کردن یک دنباله که میتواند پیاده سازی متفاوتی در برنامه عامل داشته باشد.

ب (بله . چون برنامه عامل در واقع پیاده سازی داخلی تابع عامل است.

ج (خیر . معماری عامل تنها بر روی کیفیت عامل تاثیر گذار است نه بر روی رفتار عامل در نتیجه بر روی تابع عامل نیز تاثیر گذار نیست زیرا تابع عامل بیانگر رفتار عامل است .

تمرین سوم :

الف (لیگ شبیه سازی فوتبال دو بعدی مسابقات آزاد ایران هر ساله در تهران برگزار میشود این مسابقات در سطح بین المللی انجام میگردد .

در لیگ شبیه سازی فوتبال دوبعدی، دو تیم متشکل از ۱۱ بازیکن مجازی در یک محیط شبیه سازی شده با یکدیگر بازی می کنند. تلاش شده است که در این محیط شبیه سازی شده، بسیاری از اعمال و حسگرهای روبات های واقعی و شرایط واقعی بازی شبیه سازی شود.

هر بازیکن مجازی به طور مجزا از بازیکن های دیگر، اطلاعاتی را برای سرور شبیه ساز ارسال و از آن دریافت می کند. اطلاعاتی که هر بازیکن دریافت می کند شامل اطلاعات محیطی برای آن بازیکن مانند فاصله ی بازیکن تا نقاط مختلف زمین و مکان نسبی توپ نسبت به بازیکن است. همچنین هر بازیکن اطلاعاتی مانند چگونه حرکت کردن و چگونه شوت زدن را برای سرور شبیه ساز می فرستد. اطلاعاتی که بازیکن ها دریافت می کنند، دارای خطا هستند.

در این مسابقات دو تیم از کشور چین و چند تیم از ایران حضور پیدا کردند.

ب (

در این مسابقه هر بازیکن یک عامل است و محیط آن زمین مسابقه ای ایست که عامل های دیگر نیز در آن حضور دارند . معیار کارایی در این مسابقه تعداد گل بیشتر زدن و خوردن گل کمتر برای هر تیم است در واقع هر حرکت بازیکن باید در راستای به ثمر رساندن گل بیشتر و خوردن گل کمتر باشد. محیط مسابقه زمین مسابقه میباشد و عملگر ها برای یک بازیکن توابع و دستورات حرکتی و یا ضربه به توپ میباشد همچنین حسگر هر عامل توابع دید او میباشد که میتواند یک محدوده به خصوصی را ببیند . محیط در این مسابقه جزیی قابل مشاهده است ، چند عامله است ، با وجود عوامل طبیعی در مسابقات مانند باد میتواند تصادفی باشد همچنین ترتیبی ست و پویا ست چون با گذشت زمان دایما زمین مسابقه تغییر میکند و پیوسته نیز میباشد.

ج (یک عامل برای این مسابقه طراحی میکنیم :

در این طراحی از معماری عامل های یادگیرنده استفاده میکنیم :

این معماری چهار بخش اصلی دارد : ۱- performance element - ۲ critic - ۳ learning element - ۴ problem generator

در این مساله نیازی به problem generator نخواهیم داشت زیرا عامل در طول بازی مسایل مختلفی برایش ایجاد میشود . حال به توضیح بخش های مختلف میپردازیم :

۱- Performance element

در این بخش توابعی وجود دارند که action های مربوط به ربات را میتواند انجام دهد . مثلا کارهایی مثل حرکت کردن ، شوت کردن توپ ، با توپ حرکت کردن و ورودی این بخش میتواند از learning element بیاید برای اینکه از تصمیمات و تجربه های قبلی نیز استفاده بشود .

۲- Critic

این بخش به نقد کار های عامل می پردازد . مثلا هنگام شوت کردن اگر نتوانست خوب شوت کند آن را تنبیه میکند یا اگر حرکتی کرد که منجر به یک گل شد میتواند آن حرکت را تشویق کند . نتایج این تشویق ها و تنبیه ها باید تحویل بخش learning element شود تا آن بخش بتواند تصمیم گیری داشته باشد .

۳- Learning element

با توجه به دریافت هایی که از بخش critic دارد میتواند دانش خود را بالا ببرد و این اطلاعات را به بخش performance انتقال میدهد . مثلا وقتی موقع شوت کردن عامل تشویق شد و این نتیجه به این بخش منتقل گشت این تجربه را به performance element منتقل میکند و از این به بعد هرگاه عامل بخواهد شوت کند با نظر به این تجربه سعی میکند شوت کند تا نتیجه بهتری بدست آید.

در این عامل internal state شامل وضعیت مکانی ما بقی اعضای تیم خودی و تیم حریف و مکان توپ است برای به روزرسانی این internal state سه راه پیشنهاد میشود :

۱- خود عامل از دریافتی سنسور هایش استفاده کند و آنچه که در محدوده دیدش میباشد را به روز رسانی کند . مثلا وقتی

در محدوده دید او مکان برخی از عامل ها مشخص است باید آن ها را به روزرسانی کند .

۲- با استفاده از قوانین فیزیکی محیط آن را به روز رسانی کند مثلا وقتی توپی را شوت میکند اگر هم از محدوده دید او خارج

شد بتواند با استفاده از قوانین مکان آن را حدس بزند و به روز رسانی کند

۳- همچنین با استفاده از منطق طبیعی بتواند این کار را بکند مثلا وقتی توپ در اختیار تیم خودی است و به سمت دروازه حریف

حمله میشود پس باید بازیکن های خودی به سمت جلو حرکت کنند در این صورت میتواند مکان آن ها را حدس بزند و به

روز رسانی نماید

دریافت های این عامل شامل همه آن چیزی ست که در محدوده دید او میباشد مانند مکان عامل های دیگر خودی یا تیم حریف

و همچنین مکان توپ . همچنین عمل های او شامل شوت کردن یا نگاه کردن و چرخیدن بدون توپ حرکت کردن می باشد .

تمرین چهارم :

الف (استراتژی عقلانی برای بازی این است که همواره گزینه B را انتخاب کنیم زیرا یا ۵ امتیاز از حریف جلو می‌گیریم و یا مساوی می‌کنیم . در این صورت امتیاز مورد انتظار ۳ است زیرا در نیمی از اوقات (حریف تصادفی است) ۵ و در نیمی امتیاز ۱ می‌آورد .

ب (عامل عقل کل دقیقاً در جایی که عامل تصادفی a می‌آورد b را انتخاب می‌کند و در جایی که عامل تصادفی b را انتخاب کند باز هم b را انتخاب می‌کند که بتواند ۱ امتیاز بگیرد و به صورت میانگین امتیاز ۳ خواهد بود .

عامل عقل کل همواره بهترین نتیجه و بیشترین امتیاز ممکن را بدست می‌آورد یعنی بیشتر از آن دیگر امکان پذیر نخواهد بود . در این مثال عامل عقلایی نیز الگوریتمی دارد که طبق آن چه گفته شد سود موردانتظار برابر سود واقعی و قطعی می‌باشد .