تکنیک "بهترین گوشه در مربع حالت" برای عمومی سازی حالات محیطی در یک دامنه چند عامله همکاری گرا

محمدرضا خجسته محمدرضا میبدی

آزمایشگاه سیستمهای نرم افزاری

دانشکدهٔ مهندسی کامپیوتر

دانشگاه صنعتی امیرکبیر

تهران ایران

email : mrkhojas@ce.aku.ac.ir

meybodi@ce.aku.ac.ir

چکیده

عاملها موجودات نرم افزاری هستند که بطور پیوسته و خود مختار و در محیط بخصوصی کار می کنند (که مخصوص کار در آن محیط طراحی شده اند). برای عاملها نیازهایی از قبیل واکنشی بودن نسبت به کنش های محیطی، خود مختاری در انتخاب مسیر و ادامه آن، قابلیت تطبیق و یادگیری و ... ضروری به نظر می رسد. امروزه مطالعه سیستمهای مبتنی برعاملها به یک موضوع مهم آکادمیک تبدیل شده است که کاربردهای تجاری و صنعتی فراوانی را نیز دربر دارد. در سیستمهای چندعامله، چندین عامل هوشمند با قابلیت برقراری ارتباط با یکدیگر، جهت رسیدن به مجموعه ای از اهداف، با هم همکاری می کنند. بدلیل پیچیدگیهای موجود در محیطهای چندعامله پویا و متغیر نیاز به روشهای یادگیری ماشین در چنین محیطهایی بخوبی احساس می شود.همچنین بدلیل وجود تعداد حالات بسیار زیاد در دامنه های چند عامله پیچیده، اختیار روشی برای عمومی سازی حالات محیطی، امری ضروری است که انتخاب مناسب چنین روشی، در تعیین حالات و اعمال عامل نقشی تعیین کننده دارد. در این مقاله به معرفی و پیاده سازی تکنیک "بهترین گوشه در مربع حالت" پرداخته شده است که در آن سعی شده است که فضای حالات پیوسته و بسیار وسیع عامل به فضای حالات گسسته و محدود نگاشته شود که در این راه از مشبک ساختن محیط دامنه عامل سود جسته ایم. در این مقاله با استفاده از بستر تست شبیه سازی فوتبال روباتها و با اختیار اتوماتای یادگیر بعنوان مدلی برای یادگیری عاملهای عضو یک تیم، به بررسی کارآیی این تکنیک در عمومی

سازی حالات محیطی در یک دامنه چند عامله همکاری گرا پرداخته شده است.

کلمات کلیدی: اتوماتای یادگیر، عامل، سیستمهای چندعامله، فوتبال شبیه سازی شده روباتها، تیم عاملها، همکاری

۱. مقدمه

عامل، موجود خودمختاری است که ویژگیهایی از قبیل اجتماعی بودن، واکنشی بودن و پیش فعال بودن را دارا می باشد. عاملها در جهت دست یافتن به اهداف بهتر خود یا جامعه ای که در آن زندگی می کنند با یکدیگر همکاری می کنند. ارتباطات می توانند عاملها را قادر سازد که عمل ها و رفتارهای خود را هماهنگ سازند[۱]. وقتی که یک گروه از عاملها در یک سیستم چند عامله در یک هدف [illegible] از مدت سهیم باشند، آنها در واقع تشکیل یک تیم را می دهند. اعضای یک تیم (یا هم تیمی ها) رفتار[1] خود را براساس سازگارکردن فرآیندهای شناختی[2] خود و نیز براساس تأثیر مستقیم بر ورودیهای قابل دریافت همدیگر (از طریق فعالیتهای ارتباطی)، با همدیگر هماهنگ[3] می کنند. دیگرعاملهای موجود در محیط نیز که اهدافی متقابل با هدف درازمدت این تیم دارند، حریفان یا دشمنان[4] این تیم محسوب می شوند. در واقع، یک گروه از عاملها ، تشکیل یک تیم همکاری کننده[5] را می دهند وقتی که :

- همه عاملها یک هدف مشترک دارند.
- هر عامل باید سهم خود را در رسیدن به هدف مشترک گروه انجام دهد.
- هر عامل به یک تقاضا جهت گرفتن سهم خود جواب می دهد.

بعنوان بستری برای تست و پیاده سازی سیستمهای چند عامله می توان به فوتبال روباتها یا روبوکاپ[۶] اشاره کرد. فوتبال روباتها مثالی از یک محیط پیچیده است که در آن چند عامل باید جهت رسیدن به اهداف تیمی، با هم همکاری کنند. فرآیندهای رفتاری و تصمیم گیری می توانند از ساده ترین رفتارها، همانند حرکت مستقیم به طرف توپ تا پیچیده ترین استدلال ها که استراتژی های تیم خود و تیم مقابل را روشن می سازند، تشکیل شوند.

به عنوان مثالهایی در رفتارهای چند عامله در این بستر، می توان به این موارد اشاره کرد که یک روبات مشخص، چه موقع باید به سمت توپ حرکت کند یا کار دیگری را انجام دهد و یا این که وقتی که یک روبات مشخص توپ را در اختیار دارد، آیا باید حرکت کند، پاس دهد و یا این که عمل شوت را انجام دهد (و در اینصورت با چه سرعتی ؟) تا کل تیم را به سمت یک هدف مشخص دسته جمعی (که در این حالت گل زدن است) سوق دهد.

بدلیل وجود پیچیدگیهای موجود در بسترهایی مانند بستر شبیه سازی فوتبال روباتها و رویارویی عاملهای موجود در چنین محیطهایی با حالات بسیار زیاد، متنوع و متغیر، ناگزیر به استفاده از روشهای یادگیری ماشین می باشیم. نیز بدلیل وجود تعداد حالات بسیار زیاد در دامنه های چند عامله پیچیده، اختیار روشی برای عمومی سازی حالات محیطی، امری ضروری است که انتخاب مناسب چنین روشی، در تعیین حالات و اعمال عامل نقشی تعیین کننده دارد.

در این مقاله به معرفی و پیاده سازی "تکنیک بهترین گوشه در مربع حالت" پرداخته شده است که در آن سعی شده است که فضای حالات پیوسته و بسیار وسیع عامل به فضای حالاتی گسسته و محدود نگاشته شود که در این راه از مشبک ساختن محیط دامنه عامل سود جسته ایم. ضمن آنکه به شکل ظریفی نیز سعی شده است که مواردی چون فاصله و زاویه که مفاهیمی پیوسته هستند در تعیین حالت عامل (نسبت به اطراف خود و در محیط خود) در این تکنیک نقش داشته باشند.لازم بذکر است که ما از اتوماتاهای یادگیر بعنوان مدلی برای یادگیری، در همکاری بین عاملهای عضو یک تیم استفاده کرده ایم [۱۲].

در این راستا سعی شده است که با پیاده سازی شبیه سازیهایی از تیمهای ۲ نفره، ۵ نفره، و ۱۱ نفره از عاملهای مجهز به اتوماتای یادگیر و مقایسه آن با تیم مشابه و بدون خاصیت یادگیری در بستر شبیه سازی فوتبال روباتها، در دو حالت عمومی سازی ساده حالات محیطی و عمومی سازی حالات محیطی با استفاده از تکنیک بهترین گوشه در مربع حالت، کارآیی این تکنیک در عمومی سازی حالات محیطی در یک دامنه چند عامله همکاری گرا، مورد ارزیابی قرار بگیرد.

در ادامه، ابتدا توصیفی از بستر تست شبیه سازی فوتبال آمده است و در ادامه، تکنیک بهترین گوشه در مربع حالت، بعنوان روشی برای عمومی سازی حالات محیطی، شبیه سازیها و نتایج خود را ارائه می دهیم.

۲. کارگزار فوتبال روباتها[۷]

فوتبال روباتها مثالی از محیط و وظایف پیچیده است که عاملهایی چند باید جهت رسیدن به اهداف تیمی ، با هم همکاری کنند. فرآیندهای رفتاری و تصمیم گیری می توانند از ساده ترین رفتارها ، همانند حرکت مستقیم به طرف توپ تا پیچیده ترین استدلال ها که استراتژی های تیم خود و تیم مقابل را روشن می سازند، تشکیل شوند.

سیستمهای چند عامله در محیط های پیچیده و زمان واقعی[۸] نیاز به عاملهایی دارند که بتوانند بطور مؤثر هم بعنوان یک عامل خودمختار[۹] عمل کنند و هم بعنوان عضوی از یک تیم. تأکید ما در این مقاله بر روی سیستمهایی متشکل از چند عامل خودمختار است که می توانند در محیطهای زمان واقعی، نویزی[۱۰]، نیاز به همکاری[۱۱] و دارای دشمن با اهداف متقابل[۱۲] عمل کنند.

به دلیل پیچیدگی ذاتی این نوع از سیستمهای چند عامله، یادگیری ماشین روشی جالب برای ترکیب با آن به شمار می رود. درحقیقت, یادگیری این قابلیت را دارد که عاملها را به تعداد زیادی فعالیت در رده های مختلف مجهز کند. این رده ها می توانند از رده رفتارهای فردی در یک تیم، تا رده رفتارهای همکاری و سطح بالای مشترک تیمی متغیر باشند.

ذیلا چهار خاصیت از دامنه های پیچیده سیستمهای چند عامله را توضیح می دهیم :

زمان واقعی - محیطها و دامنه هایی را زمان واقعی گویند که در آنها موفقیت بستگی به عمل کردن مناسب و در پاسخ یک محیط متغیر دینامیک است.

نویزی - دامنه هایی را نویزی می گویند که در آنها عاملها نمی توانند بطور صحیح و دقیق دنیا را درک کنند و نیز نمی توانند بطور دقیق برروی آن تأثیر بگذارند.

همکاری گرا - دامنه هایی را گویند که در آنها یک گروه از عاملها در یک هدف مشترک سهیم هستند.

دارای دشمن - به دامنه هایی می گویند که در آنها عاملهایی با اهداف رقابتی وجود دارند.

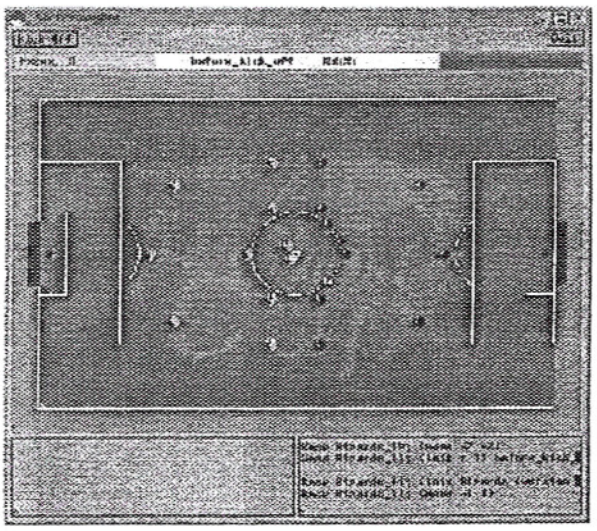
فوتبال روباتها، دامنه ای است که با خاصیت های در نظر گرفته شده فوق برای دامنهٔ کار، تناسب کامل دارد، در دسترس است و نیز به اندازهٔ کافی پیچیدگی دارد. همانگونه که هدف اساسی هر بستر تستی این است که این امکان را برای ما ایجاد کند که بتوانیم ایده های خود را در آن مورد ارزیابی قرار دهیم، فوتبال روباتها نیز این امکان را به ما می دهد که ایده های وجود در دنیای واقعی را در آن مورد پیاده سازی قرار دهیم و از این جهت یک بستر تست عالی برای پروژه هایی از این دست به شمار می رود [۲]. ما در این تحقیق از فوتبال شبیه سازی شده روبوتیک استفاده کرده ایم.

در دامنه هایی با چنین خواص پیچیده، عاملها قادر نخواهند بود که بطور مؤثر بیاموزند که چگونه از حسگرهای خود یک نگاشت مستقیم به محرک های خود ایجاد کنند، حتی اگر حالتهای گذشته دنیای خود را ذخیره کرده باشند [۲].از آنجا که ما از کارگزار شبیه ساز فوتبال روباتها بعنوان بستر تست خود استفاده کرده ایم، ذیلا به شرح خلاصه ای از این بستر می پردازیم.

بسترفوتبال شبیه سازی شده روبوتیک، محیطی را فراهم می کند که کاملا"توزیع شده می باشد. در این محیط حالت مخفی[13] وجود دارد، به این معنی که هر عامل تنها یک دید جزیی از دنیا را (در هر لحظه) دارا است. هم چنین عاملها دارای حسگرهای نویزی[14] و نیز محرک های نویزی هستند، به این معنی که آنها نمی توانند دنیا را به شکل دقیق درک کنند و نیز نمی توانند بر روی آن نیز، دقیقا" به همان شکل می خواهند تأثیر بگذارند. بعلاوه، سیکل ها[15] و زمان های دریافت[16] و عمل[17] به صورت آسنکرون هستند که این خود مانعی برای این مورد است که از کاربردهای سنتی هوش مصنوعی که در آنها از ورودی دریافتی برای تعیین عمل استفاده می شود ، استفاده کرد. در این محیط، فرصتهای ارتباطی محدود هستند، عاملها باید بصورت زمان واقعی تصمیم گیری نمایند و نیز اعمال انجام شده توسط دیگر عاملها، چه هم تیمی و چه دشمن و نیز نتایج تغییر وضعیت حاصل از این اعمال برای هر عامل بخصوص ناشناخته[18] هستند.

کارگزار فوتبال روباتها [3] بعنوان پایه ای برای مسابقات بین المللی موفق [4] و نیز چالشهای تحقیقاتی [5] بکار رفته است. کارگزار فوتبال یک دامنه واقعیت گرا[19] و پیچیده است. برخلاف بسیاری از دامنه های هوش مصنوعی، محیط کارگزار فوتبال سعی نموده است که حداکثر ممکن پیچیدگیهای دنیای واقعی را درون خود مورد شبیه سازی قرار دهد. این کارگزار یک سیستم روباتیک فرضی را با ترکیب خصوصیاتی از سیستم های برنامه ریزی شده[20] و موجود[21] و نیز بازیکنان فوتبال انسانی پیاده سازی می نماید.

خصوصیت نویزی بودن حسگر و محرک در کارگزار از سیستم های روباتیک واقعی برداشت شده است، حال آن که خیلی از خصوصیات دیگر، از جمله قوت[22] محدود و نیز دید محدود، از پارامترهای انسانی برداشت شده اند. جزئیات موجود، مفاهیم و پارامترهای مهم در گزارش کارگزار فوتبال [6] آمده است. محیط شبیه ساز از یک ابزار تصویری بصورتی که در شکل ۱ نشان داده شده است استفاده می کند.



شکل ۱- محیط شبیه سازی فوتبال روباتها

محیط شبیه ساز فوتبال از یک مدل عمل گسسته[23] استفاده می کند، بدین صورت که اعمال بازیکن را در طی یک سیکل شبیه سازی[24] بطول ۱۰۰ میلی ثانیه مورد جمع آوری قرار می دهد، ولی تنها در پایان سیکل، آنها را اجرا می کند و درحقیقت دنیا را به روز می سازد. اگر یک کارفرما ، بیش از یک فرمان اجرایی را در یک سیکل شبیه سازی بفرستد، کارگزار یکی از آنها را بصورت تصادفی برای اجرا انتخاب می کند. بنابراین بجاست که هر کارفرما در حین هر سیکل شبیه سازی حداکثر یک فرمان را به کارگزار بفرستد. از سویی دیگر، اگر یک کارفرما در حین یک سیکل شبیه سازی هیچ فرمان اجرایی را به کارگزار نفرستد، فرصت عمل را در حین آن سیکل از دست می دهد که می تواند در یک محیط زمان واقعی و دارای دشمن ضرر زیادی باشد، زیرا در همین حین که عامل بیکار بوده است، حریفان ممکن است این فرصت را بدست آورند که برتر ظاهر شوند. در هر سیکل، سیبه ساز یک واحد به شمارندهٔ شبیه ساز اضافه می کند.

حس کردن و عمل نمودن بصورت آسنکرون، مخصوصا" وقتی که حس کردن می تواند در فواصل زمانی غیر قابل پیش بینی اتفاق بیفتد، برای عاملها موردی بسیار چالش برانگیز است. در حقیقت عاملها باید بین نیاز به عمل کردن بصورت منظم و سریع و نیز نیاز به جمع آوری اطلاعات از محیط، یک تعادل برقرار کنند. باید اضافه کرد که این مورد، یعنی حس کردن و عمل کردن آسنکرون، تنها یکی از موارد پیچیدگی دنیای واقعی است که شبیه ساز آن را پوشش می دهد.

کارگزار فوتبال یک شبیه ساز دو بعدی است. حداکثر ۲۲ بازیکن و یک توپ در هر لحظه می توانند در زمین وجود داشته باشند که همهٔ آنها بصورت دایره هایی مدل می شوند. همچنین چندین نشانه[25] قابل رؤیت در زمین و پیرامون آن وجود دارند، شامل پرچمها و خطهای کناری زمین که در اطراف زمین وجود دارند.

بازیکنان ز توپ همگی اشیاء متحرک به شمار می روند. در سیکل شبیه ساز برای زمان t ، هر شیء در یک موقعیت مشخص (p_x^t , p_y^t) قرار دارد و سرعتی برابر با (v_x^t , v_y^t) نیز دارا است . همچنین هر بازیکن در لحظهٔ t به یک جهت مشخص θ^t رو کرده است. این موقعیت ها بصورت داخلی بصورت اعداد اعشاری نگهداری می شوند، ولی حسهایی که برای بازیکنان توسط کارگزار فرستاده می شود حداکثر دارای دقتی با یک رقم اعشارمی باشند. بنابراین فضای حالت دریافت شده با اغلب پارامترهای کارگزاری که ما استفاده می کنیم بیش از 10^{22} یا 10^{198} حالت است. بدینصورت که هرکدام از این ۲۲ بازیکن میتوانند در یکی از ۳۶۰۰ ×۱۰۵۰ ×۶۸۰ مکان باشند. با درنظرگرفتن توپ، سرعتها و حالتهای گذشته، فضای حالات واقعی حتی بسیار بزرگتر از این مقدار می باشد [2].

کارگزار فوتبال، نویز با توزیع احتمال یکنواختی به حرکات همه اشیاء اضافه می کند. نیز کارگزار فوتبال مانع از آن می شود که یک بازیکن بتواند دائما" با حداکثر سرعت بدود و این کار را با انتساب یک قوت محدود به هر بازیکن انجام می دهد.

عاملها سه نوع اطلاعات درکی و حسی متفاوت از کارگزار دریافت می کنند: شنوایی[۲۶] ، بینایی[۲۷] و فیزیکی[۲۸][۶]. ارتباطات در کارگزار فوتبال بدینصورت مدل شده است که از محیطی نویزی[۲۹] و با پهنای باند کم استفاده شده است. در واقع همه ۲۲ عامل (هرتیم ۱۱ عامل) از یک کانال ارتباطی غیر مطمئن استفاده می کنند.

لازم بذکر است که ما در این تحقیق، از مدل اتوماتای یادگیر برای یادگیری عاملهای خود بجهت عمل در این دامنه پیچیده استفاده کرده ایم [۱۲].

اتوماتاهای یادگیر در محیطی تصادفی عمل نموده و قادر هستند که بر اساس ورودیهای دریافت شده از محیط، احتمال انجام عملیات خود را بروز در آورده تا بتوانند از این طریق کارآیی خود را بهبود بخشند. اتوماتای یادگیر یک مدل انتزاعی است که تعداد محدودی عمل را می تواند انجام دهد. هر عمل انتخاب شده توسط محیطی احتمالی ارزیابی می گردد و پاسخی به اتوماتای یادگیر داده می شود. اتوماتای یادگیر از این پاسخ استفاده نموده و عمل خود برای مرحله بعد انتخاب می کند [۷]، [۸]. برای مطالعه بیشتر درباره اتوماتا های یادگیر می توان به [۷]، [۸] ،[۹]، [۱۰]، و [۱۱] مراجعه نمود.

۳. یک عمومی سازی ساده از حالات محیطی برای عاملها

بستر انتخابی ما همانگونه که شرح آن رفت، محیط شبیه ساز فوتبال روباتها می باشد که با توجه به خصوصیات فوق الذکر، کلیه خصوصیات لازم جهت ایجاد یک محیط شبیه سازی پیچیده چند عامله را در اختیار ما قرار می دهد.

لازم بذکر است که ما در کلیه شبیه سازیهای خود در این تحقیق، از امکان شنوایی بازیکنان بسیار کم استفاده کرده ایم و بر روی شنوایی محدود بازیکنان تکیه نداشته ایم.

همچنین بدلیل وجود مواردی چون مشکلات انتشار به عقب و نیزحالات مخفی بسیار زیاد در این دامنه، ترجیح داده شده است که پاداش یا جزا بصورت محلی (و تنها به آخرین بازیکن عمل کننده) داده شود و در واقع هر عمل بازیکن تا به سر انجام رسیدن همان عمل ارزیابی می شود. در واقع در دامنه ای با چنین پیچیدگی، شاید این مورد، مناسب نیز تلقی می شود.

در سری اول شبیه سازیهای انجام شده توسط ما، دو بازیکن فوتبال با خاصیت یادگیری از نوع اتوماتای یادگیر را در مقابل یک و دو بازیکن بدون خاصیت یادگیری قرار دادیم و تاثیر این نوع یادگیری را در آنها ارزیابی کردیم[۱۲]. با توجه به تعداد حالات بسیار زیاد در دامنه فوتبال شبیه سازی، سعی کردیم که کلیه حالات محیط را درون ۴ حالت برای بازیکن دارای توپ و ۴ حالت برای بازیکن بدون توپ خلاصه کنیم [۱۳].

برای هریک از این چهار حالت نیز، یک اتوماتای یادگیر با ساختار ثابت و با عمق حافظه ۳ درنظر گرفته شد که هر اتوماتا، یکی از دو عمل تعریف شده (پاس به هم تیمی و یا شوت به طرف دروازه حریف) را در هر لحظه مورد انتخاب و اجرا قرار می داد.

همانگونه که نتایج نشان دادند[۱۳]، تیم دارای اتوماتای یادگیری در قیاس با یک تیم بدون یادگیری بسرعت یاد می گیرد که در چه حالاتی، باید چه اعمالی را انجام دهد و بدین لحاظ براحتی بر حریف خود غلبه می کند. نیز روش یادگیری همزمان ما این حسن را دارد که به تیم ما اجازه می دهد خود را با نحوه بازی تیم حریف تا حد زیادی وفق دهد.

اما برای یک یادگیری بهتر، بنظر می رسید که باید عمومی سازی بهتری از محیط انجام دهیم تا حالات مشابه بشتری به یک حالت واحد نگاشت شوند. در واقع، ایرادی که در تعیین حالت بازیکن صاحب توپ در شبیه سازیهای سری قبل موجود بود، این بود که امکان نگاشت حالات بسیار متفاوتی از دنیای پیرامون عامل صاحب توپ به یک حالت واحد وجود داشت و لهذا ممکن بود عملی که عامل بعنوان عمل صحیح در آن حالت خود انتخاب کرده بود، در واقع، بهترین عمل ممکن برای آن نبوده باشد. چه در واقع، حالت عامل بدرستی تشخیص داده نشده بود.

بدین صورت، عامل در شناخت حالت خود دچار مشکل است و لهذا نیاز به تعداد حالات بیشتر و یا "عمومی سازی" بهتری از حالات محیطی برای عامل کاملا احساس می شود. ایراد دیگری نیز که در رابطه با امکان عمل برای عامل صاحب توپ (و حتی بدون توپ) در شبیه سازیهای فصل قبل موجود بود، وجود تعداد کم اعمال ممکن برای عامل (که تنها در پاس دادن به هم تیمی و یا شوت به دروازه حریف خلاصه می شد) بود.

لهذا روش دیگری را برای عمومی سازی بهتر حالات محیطی در دامنه های پیچیده را مورد طراحی و پیاده سازی قرار دادیم.در این روش که ما آن را "تکنیک بهترین گوشه در مربع حالت" نامیده ایم، عمومی سازی مناسبی از حالات محیطی نسبت به شبیه سازیهای قبلی صورت گرفته است.

ما شبیه سازیهای قبلی (که با اتوماتاهای با ساختار ثابت انجام شده بودند) بعلاوه شبیه سازیهای دیگری از عاملهای مبتنی بر اتوماتاهای یادگیر با ساختار متغیر و یادگیری Q را نیز با استفاده از این تکنیک برای "تعیین حالت عامل" مورد مقایسه و پیاده سازی قرار دادیم[۱۳].نتایج نشان دادند [۱۳] که با این تکنیک، عامل در تعیین حالت خود بهتر عمل می کند و در نتیجه بهتر می تواند عمل مناسب خود در هر حالت را نیز انتخاب نماید.

شاید شاخص ترین مورد در این تکنیک، نگاشت فضای حالات پیوسته و بسیار وسیع به فضای حالاتی گسسته و محدود (بامشبک ساختن محیط دامنه عامل) برای عامل می باشد. ضمن آن که به شکل ظریفی نیز سعی شده است که مواردی چون فاصله و زاویه که مفاهیمی پیوسته هستند در تعیین حالت عامل (نسبت به اطراف خود و در محیط خود) در این تکنیک نقش داشته باشند. در این تکنیک، برای هر عامل ۸ حالت (و لهذا ۸ اتوماتا) و برای هر اتوماتا، ۸ عمل ممکن در نظر گرفته شده است که هرکدام از این اعمال ارسال توپ به یکی از ۸ جهت اصلی اطراف مربع حالت عامل، برحسب حالت و مورد است.

ایدهٔ جدید دیگری که در ادامه تکنیک فوق در [۱۳] مطرح شده است، پیاده سازی تیمی با نام "یادگیری مربعهای کامل" است که در آن در هر مربع حالت از

فضای مشبک محیط عامل، ۸ اتوماتا تعبیه کرده ایم و بدین صورت، عامل، اتوماتاهای مورد نیاز خود را با حرکت در محیط، با خود جابجا نمی کند.

ضمن آن که تیمی با عنوان "نگاشت ثابت" نیز که در واقع یک تیم کد شده بهینه با دست است در این راستا ایجاد شده و با تیمهای یادگیر مورد مقایسه قرار گرفته است. نکته مهم دیگر تکنیک "بهترین گوشه در مربع حالت"، تصمیم گیری عامل بر روی فضای محدود و "محلی" خود در جهت همکاری با دیگر عاملها برای برآوردن هدف "سرتاسری" تیم است.

٤. تکنیک بهترین گوشه در مربع حالت

در واقع در دومین سری شبیه سازیهایمان، سعی شد که با رفع ایرادهای فوق الذکر و با ایجاد مدل دنیای واقعی تر برای عامل، به همکاری بهتری در بین عاملهای موجود در یک تیم دست یابیم.

همانگونه که در فصلهای قبلی نیز توضیح داده شد، تعداد حالات موجود در دامنه فوتبال روبوتیک شبیه سازی شده، بسیار زیاد است و لهذا امکان در نظر گرفتن کلیه این حالات برای یک عامل، غیرممکن بنظر می رسد. در واقع مهمترین کار در این قسمت، یک عمومی سازی مناسب از حالات برای عامل است. با فرض آنکه مجموعه حالات دامنه را S بنامیم و V، مجموعه حالات نگاشت باشند، در واقع می خواهیم تابع زیر را بیابیم:

$$f: S \rightarrow V$$

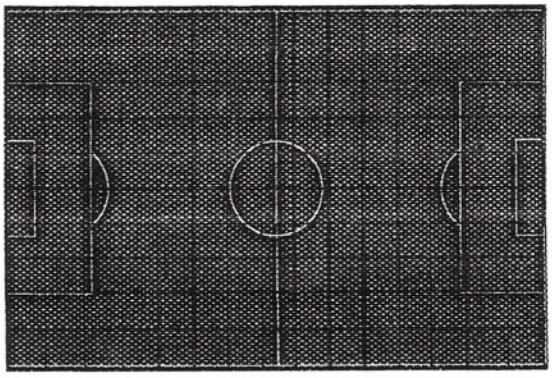
بدینصورت، پس از در داست داشتن f، عامل می تواند از V برای یادگیری و عمل خود در محیط استفاده نماید.

غیر از این تابع، یک تابع تقسیم وظایف P نیز موجود است که در واقع، اصل همکاری را می رساند و مجموعه حالات دامنه را بین عاملهای واقع در محیط تقسیم بندی می کند:

$$P: S \rightarrow M$$

تابع فوق، در واقع، فضای حالت را به $|M|$ بخش مجزا تقسیم می کند و هر بخش به حداقل یک عامل (جهت یادگیری و عمل در آن بخش) سپرده می شود. با توجه به توضیحات بالا و با فرض این که مجموعه اعمال عامل، A باشد، عامل در هر کدام از $|V|$ حالت، دارای $|A|$ عمل ممکن خواهد بود و بدین ترتیب، مجموعه مورد یادگیری عامل حداکثر شامل $|V| \times |A|$ خواهد بود که این مورد با انتخاب معقول مجموعه های V و A، امکان یادگیری مناسب با مثالهای محدود را در دامنه ای پیچیده و همزمان به ما می دهد.در واقع مجموعه های V و A باید بگونه ای انتخاب شوند که تا حد امکان، در بر گیرنده کلیه حالات و اعمال باشند و در واقع، نگاشت های خوبی از مجموعه های حالات و اعمال کلی ممکن در دامنه محیط در بردارنده عامل بحساب آیند.

ایده ما برای عمومی سازی محیط بدینصورت است که کل فضای مستطیل شکل زمین فوتبال (که در واقع محیط دامنه چندعامله ما محسوب می شود) را به ۱۵۰ مربع یکسان با عرض ۷ متر تقسیم کرده ایم و در واقع، فضای زمین بازی، به شبکه ای از مربعات مساوی تقسیم شده است (شکل ۲).در واقع، ما با این شبکه مربعها، محیط پیوسته پیرامون عامل را به محیطی گسسته تبدیل کرده ایم.



شکل ۲- تقسیم زمین به ۱۵۰ مربع ۷ × ۷ متر مربعی

بدینصورت، در هر لحظه از زمان بازی، هر عامل درون یکی از این مربعها قرار دارد. با در نظر داشتن این مورد که هر عامل دارای دیدی محدود می باشد و نمی تواند بر روی دیدی کامل از زمین (محیط) و دیگر عاملها حساب کند، ما برای هر عامل که در زمان t درون مربع حالت X قرار دارد، ۲۴ مربع اطراف آن را در همان زمان t مورد توجه قرار داده ایم. به شکل ۳ توجه فرمایید.

X-22	X-12	X-2	X+8	X+18
X-21	X-11	X-1	X+9	X+19
X-20	X-10	X	X+10	X+20
X-19	X-9	X+1	X+11	X+21
X-18	X-8	X+2	X+12	X+22

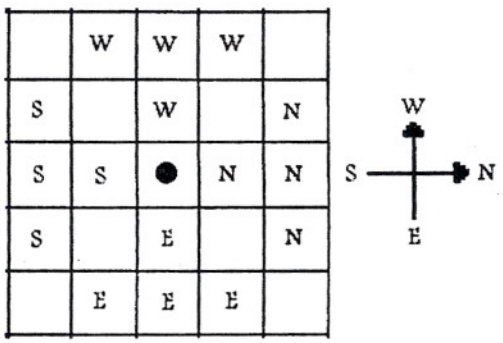
شکل ۳- ۲۴ مربع اطراف مربع دربردارنده عامل مورد نظر در مربع X

در شکل ۳، مربعهای X+10، X+11، X+1، X-9، X-10، X-11، X-1، و X+9 را مربعهای بلافاصله اطراف یا بلافاصله برای مربع مورد نظر X می نامیم. ذکر این نکته ضروری است که شماره گذاری مربعها از چپ به راست و بصورت ستونی و با شروع از شماره صفر برای اولین مربع (منتها الیه چپ و بالا) تا شماره ۱۴۹ برای آخرین مربع (منتها الیه راست و پایین) انجام گرفته است. نیز در هر ستون دارای ۱۰ مربع و (طبعا) در هر سطر دارای ۱۵ مربع می باشیم.

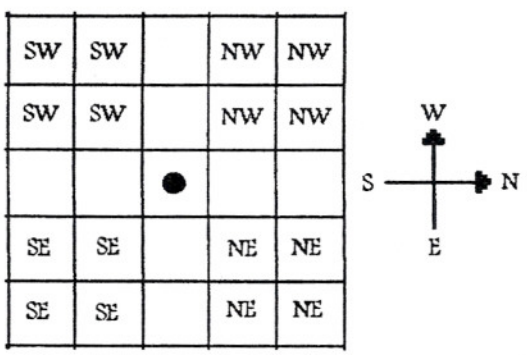
روش کار بدینصورت است که برای هر کدام از ۸ مربع اطراف عامل و در واقع برای هر یک از ۸ جهت اصلی (شمال غربی، شمال، شمال شرقی، شرق،

جنوب شرقی، جنوب، جنوب غربی، غرب) اطراف آن، یک کمیت عددی بنام "ضریب اطمینان" آن جهت در نظر گرفته شده است.

این کمیت، در واقع، نشان از ضریب اطمینانی است که عامل صاحب توپ در ارسال توپ خود به سمت مرکز (نقطه وسط) مربع بلافاصله نظیر هر جهت اصلی دارد. در شکلهای ۴ و ۵، مربعهای ۸ جهت اطراف مربع دربردارنده عامل را (با اختیار جهت شمال به سوی دروازه حریف) نشان داده ایم.



شکل ۴- مربعهای اختیار شده بعنوان مربعهای شمال غرب، شمال شرق، جنوب غرب، و جنوب شرق مربع دربردارنده عامل با توجه به موقعیت عامل مورد نظر و چهارجهت اصلی



شکل ۵- مربعهای اختیار شده بعنوان شمال، شرق، غرب، و جنوب مربع دربردارنده عامل با توجه به موقعیت عامل مورد نظر و چهارجهت اصلی

با توجه به موارد بالا و به عنوان مثال، اگر عامل ما (واقع در مربع حالت X) یکی از عاملهای حریف را در درون یکی از مربعهای شمال غرب خود (و جزء ۲۴ مربع = مربعهای X+9، X+8، X+18، و X+19) ببیند، با فرض وارد شدن یک بردار نیروی دفعی از طرف عامل حریف به خود، موقعیت بهینه جهت ارسال توپ را، در مربع بلافاصله جنوب شرقی خود (یعنی مربع X-9) می یابد و یک مقدار عددی منفی به ضریب اطمینان جهت شمال غرب خود اضافه می کند. بعبارت بهتر، عامل اطمینان خود از ارسال توپ به سمت شمال غرب خود را کاهش می دهد (و بالطبع اطمینان خود به ارسال به جهات دیگر را افزایش می دهد).

به عنوان مثالی دیگر، اگر عامل صاحب توپ مورد نظر ما (واقع در مربع حالت X) یکی از عاملهای حریف را در درون یکی از مربعهای شمال خود (و جزء ۲۴ مربع = مربعهای X+10, X+20, X+19, X+21) ببیند، با فرض وارد شدن یک بردار نیروی دفعی از طرف عامل حریف به خود، موقعیت بهینه جهت ارسال توپ را، در مربع بلافاصله جنوب خود (یعنی مربع X-10) می یابد و یک مقدار عددی منفی به ضریب اطمینان جهت شمال خود اضافه می کند.بعبارت بهتر، عامل اطمینان خود از ارسال توپ به سمت شمال خود را کاهش می دهد. این وضعیت برای عاملهای هم تیمی برعکس است. به عنوان مثال، اگر عامل صاحب توپ مورد نظر ما (واقع در مربع حالت X) یکی از عاملهای خودی را در درون یکی از مربعهای جنوب غرب خود (و جزء ۲۴ مربع = مربعهای X-21 ویا X-22 ویا X-11 ویا X-12) ببیند، با فرض وارد شدن یک بردار نیروی جذبی از طرف عامل خودی به خود، موقعیت بهینه خود را جهت ارسال توپ به آن موقعیت، در مربع بلافاصله جنوب غرب خود (یعنی مربع X-11) می یابد و یک مقدار عددی مثبت به ضریب اطمینان جهت جنوب غرب خود اضافه می کند. بعبارت بهتر، عامل اطمینان خود از ارسال توپ به سمت جنوب غرب خود را افزایش می دهد.

لازم بذکر است که ما اندازه این افزایش منفی ویا مثبت در ضریب اطمینان هر جهت را با معکوس فاصله عامل حریف و یا عامل هم تیمی از عامل مورد نظر (و در واقع صاحب توپ) در آن جهت، متناسب کرده ایم. در واقع، عاملهای نزدیک تر (و درون فاصله محلی تا ۲۴ مربع حالت پیرامون عامل صاحب توپ) تاثیر بیشتر و عاملهای دورتر (و درون فاصله محلی تا ۲۴ مربع حالت پیرامون عامل صاحب توپ) تاثیر کمتری را بر روی ضرایب اطمینان عامل صاحب توپ خواهند داشت و عاملهای خارج از فاصله محلی (خارج از ۲۴ مربع حالت پیرامون عامل)، تاثیری در تغییر ضرایب اطمینان عامل صاحب توپ ندارند.

بدینصورت، ما عامل را وادار به یادگیری و عمل بر اساس اطلاعات محلی پیرامون خود می کنیم "و هر عامل سعی می کند که بصورت محلی خوب رفتار کند".

در ادامه روند بالا، هر عامل با محاسبه برآیندی موقعیت تمامی عاملهای هم تیمی و حریف پیرامون خود (درون ۲۴ مربع)، ۸ ضریب اطمینان خواهد داشت که این ضرایب اطمینان، از دید خود عامل، نشان از کمیت بهینه بودن ۸ مربع بلافاصله مربع دربردارنده عامل، از جهت ارسال توپ به سمت مرکز آن مربعها دارند.

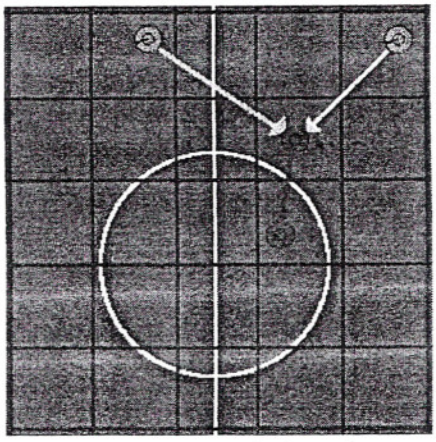
با این ترتیب، ما در واقع دارای ۸ حالت (متناظر با ۸ جهت اصلی پیرامون عامل) و ۸ عمل (متناظر با ارسال توپ به سمت مرکز ۸ مربع بلافاصله پیرامون عامل) برای عامل صاحب توپ خواهیم بود که با در نظر داشتن شعاع موقعیتی عامل در درون هر یک از مربعهای ۷×۷ متر مربعی، عمل ارسال توپ به سمت مرکز ۸ مربع پیرامون، در واقع ارسال توپ به تمامی جهات ممکن توسط عامل را در بر خواهد داشت.

لازم بذکر است که تقسیم زمین به مربعهای ۷×۷ متر مربعی بر اساس تجربه کسب شده در کار با محیط شبیه ساز و میزان جابجایی حاصله توپ در اثر هر ضربه حاصل شده است و عملا وابسته به دامنه انتخابی است. بدینصورت، فضای

حالات پیرامون عامل به نحو بسیار جالبی به ۸ حالت (در برگیرنده) کاهش می یابد و مشکل نگاشت حالات متفاوت به یک حالت بخصوص تا حد زیادی از بین می رود. بدیهی است که ممکن است در لحظه t چندین عامل هم تیمی و حریف در درون ۲۴ مربع حالت پیرامون مربع حالت عامل وجود داشته باشند. حال اگر عامل با عاملهایی درون همان مربعی قرار داشته باشند که عامل صاحب توپ در آن قرار دارد، عامل صاحب توپ بر حسب آنکه عاملهای فوق در کدام یک از ۸ جهت اصلی خود قرار دارند و باز بر طبق همان قانون جذب و دفع، ضرایب اطمینان خود را تغییر می دهد. یادآوری می شود که با نزدیک تر شدن فاصله هر عامل به عامل مورد نظر ما، تاثیر جذبی یا دفعی آن بیشتر می شود.

بدین ترتیب، اگر هرکدام از این ۸ ضرایب اطمینان (که هرکدام نشان دهنده میزان اطمینان از ارسال توپ به یکی از جهات اصلی از دید عامل صاحب توپ می باشند) را به مربعهای حالت بلافاصله اطراف عامل نظیر کنیم و هر کدام از این مربعها را نماینده یک حالت برای عامل در نظر بگیریم، هر عامل نهایتا دارای ۸ حالت خواهد بود.

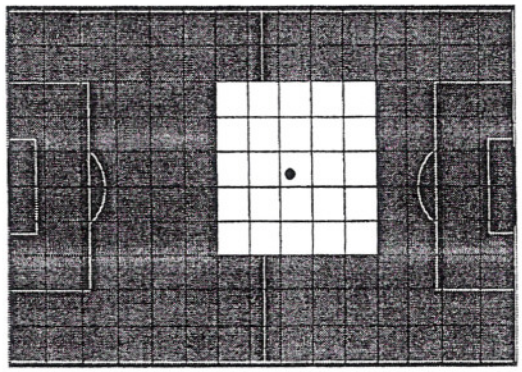
عامل صاحب توپ در هر سیکل، با توجه به برآیند جذب و دفع عاملهای هم تیمی و غیرهم تیمی بر خود، "بهترین گوشه در مربع حالت" و در واقع حالت خود را محاسبه می کند. به شکل ۶ توجه فرمایید.



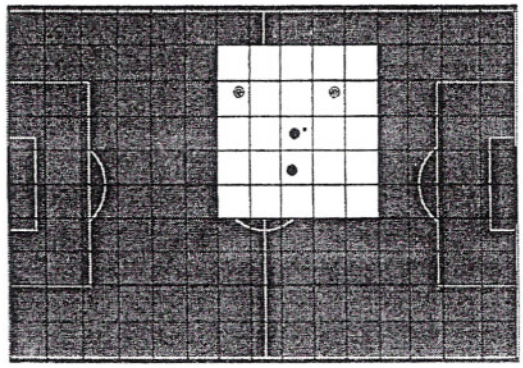
شکل ۶- نیروهای جذب و دفع فرضی به عامل صاحب توپ جهت تعیین حالت آن در روش بهترین گوشه در مربع حالت

در مورد انتخاب ۲۴ مربع اطراف برای هر عامل (که در واقع آنها را فضای محلی یا قابل اهمیت هر عامل نیز می نامیم) می توان به اشکال زیر توجه کرد.

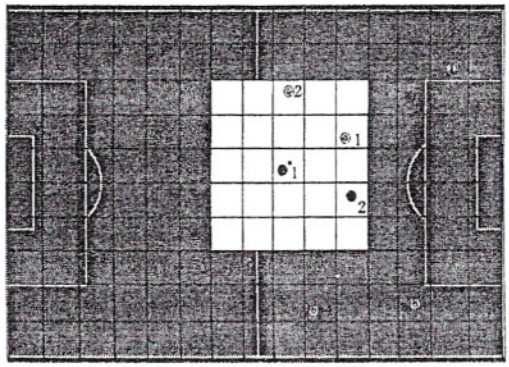
در شکل ۷، فضای محلی و قابل اهمیت برای یک عامل در روش بهترین گوشه در مربع حالت، در قیاس با کل فضای بازی نشان داده شده است. همانطور که از شکل نیز مشخص است، تکیه بر (تنها) این فضا بجای کل فضای بازی برای تصمیم گیری (مناسب محلی) یک عامل، معقول بنظر می رسد. شکل ۸، نشان می دهد که در یک بازی ۲ در مقابل ۲، عاملهای هم تیمی و حریف برای یک عامل صاحب توپ، درون این فضای محلی قرار دارند و لهذا همگی در تصمیم گیری عامل، مهم محسوب می شوند. شکل ۹ یک بازی ۴ در مقابل ۴ را به تصویر کشیده است. در این شکل، برای بازیکن صاحب توپ، بازیکن ۲ هم تیمی و بازیکنان ۱ و ۲ از تیم حریف صاحب اهمیت هستند (به این دلیل که درون فضای محلی وی قرار دارند) و سایر بازیکنان نقشی در تصمیم گیری بازیکن صاحب توپ ندارند.



شکل ۷- فضاي محلي و قابل اهميت براي يك عامل در روش بهترين گوشه در مربع حالت

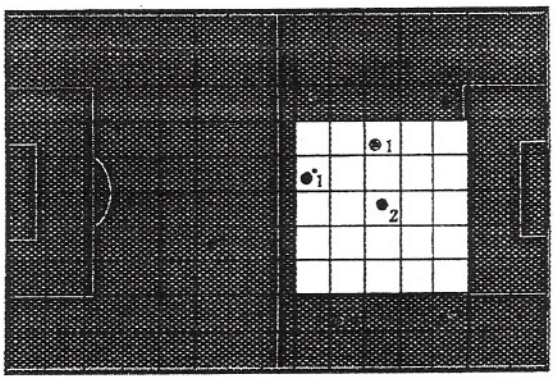


شکل ۸- فضاي محلي عامل صاحب توپ و با يكنان ه تيمي و حريف درون اين فضا در يك با ي در م ابل



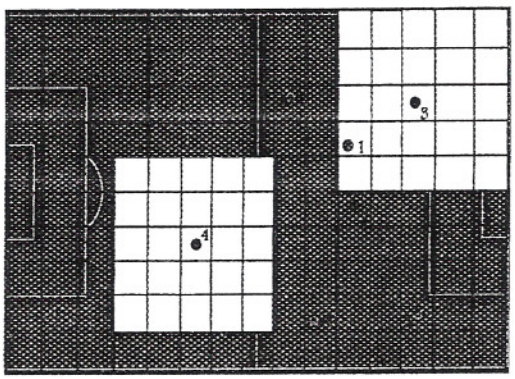
شكل - وضعيت با يكنان ه تيمي و حريف با توجه به فضاي محلي عامل صاحب توپ در يك با ي ۴ در م ابل ۴

شکل ۱۰، همان وضعیت شکل ۹ را، از دید بازیکن ۲ هم تیمی بازیکن صاحب توپ نشان می دهد.با توجه به این شکل، بازیکن ۲، تنها موقعیت عاملهای ۱ هم تیمی و ۱ حریف را در تصمیم گیری خود برای این مورد که در کجا بایستد، وارد می کند. و نهایتا شکل ۱۱ نیز وضعیت فوق الذکر را از دید عاملهای ۳ و ۴ هم تیمی (فضای حالات و قابل اهمیت برای آنها) نشان می دهد.



شکل - وضعیت با یکنان ه تیمي و حریـ با توجه به فضاي محلي عامل در یك با ي ۴ در مـ ابل ۴

با توجه به اینکه حالت موجود هر عامل در محیط در لحظه t به مقدار زیادی به فاصله و زاویه دیگر عاملها نسبت به آن بستگی دارد و با توجه به توضیحات فوق، نگاشت مناسبی از حالتهای یکسان به گوشه واحد و از حالتهای متفاوت به گوشه های متفاوت صورت می گیرد.



شکل - وضعیت با یکنان ه تیمي و حریـ با توجه به فضاي محلي عاملهاي و ۴ در یك با ي ۴ در مـ ابل ۴

ما در [۱۳] از همین مربعها برای تعیین شکل تیمی اعضای تیم و موقعیت هر کدام از اعضای تیم نیز استفاده کرده ایم.

ایده نگاشت و عمومی سازی فضای دامنه فعالیت عاملهای موجود در یک تیم به فضایی گسسته از شبکه مربعهای حالت توسط نگارنده ابداع شده است و در مراجع مربوطه از جمله [۲]، از مفهومهای پیوسته فاصله، زاویه و نظایر آن برای عمومی سازی فضای حالات عامل استفاده شده است.

نکته قابل توجه در این قسمت آن است که حالت انتخاب شده توسط عامل صاحب توپ در مرحله تعیین حالت، لزوما به این معنی نمی باشد که عامل صاحب توپ، حتما مربع بلافاصله نظیر آن حالت را جهت ارسال توپ به سمت مرکز آن، انتخاب نماید. حالت فوق الذکر، تنها نشان دهنده یکی از ۸ حالت برای عامل صاحب توپ می باشد که با توجه به توضیحات فوق، در هر حالت، عامل می تواند هرکدام از ۸ عمل ممکن خود (ارسال توپ به سمت مرکز هر یک از ۸ مربع بلافاصله اطراف خود) را مستقل از حالت انتخاب شده انتخاب نماید.

شبیه سازیهایی در [۱۳] وجود دارند که نشان می دهند هرچند تعریف ما از حالت برای عامل صاحب توپ، این ذهنیت را تداعی می کند که جهت ارسال توپ باید به سمت مرکز مربع حالت نظیر با آن حالت باشد، ولی در بسیاری از موارد، این بهترین عمل بهینه عامل ما نمی باشد. لهذا حالات ما مستقل از بهترین عمل انتخابی ما در آن حالات می باشند. بنابراین، هر عامل حداکثر دارای ۶۴ المان قابل یادگیری (۸ عمل ممکن در ۸ حالت موجود) خواهد بود. در واقع، در هر حالت، عامل (با استفاده از یک اتوماتا) می تواند یک کمیت عددی برای هر یک از ۸ عمل خود یاد بگیرد که با توجه به ماکزیموم کمیت یاد گرفته شده در هر حالت، اتوماتا، عمل بهینه خود را در آن حالت انتخاب می نماید.

سناریوی کلی برای هر عامل در شبیه سازیهای این قسمت نیز مانند قسمت قبل است با این تفاوت که اگر عامل صاحب توپ، بازیکنی غیر از خود را در مسیر به سمت دروازه حریف ببیند، با تعیین حالت خود و عمل بهینه خود در آن حالت (با استفاده از اتوماتای نظیر آن حالت)، سعی می کند حرکتی را انجام دهد که در مجموع به سود حرکت و هدف تیمی باشد.

بازیکنی که بازیکن تیم خودی را صاحب توپ ببیند نیز، موقعیت خود و دیگر بازیکنان موجود در حوزه دید خود را به بازیکن صاحب توپ اعلام می کند تا بازیکن صاحب توپ، حتی المقدور، بهترین تخمین را از حالتی که در آن قرار دارد داشته باشد. این امر با توجه به این نکته که دید هر بازیکن محدود است و هر بازیکن، کنار و پشت خود را نمی بیند، برای بازیکن صاحب توپ حیاتی است و لهذا در این مورد (و تنها در این مورد) از امکان شنوایی (و آنهم بصورت محدود) استفاده شده است. در واقع در چنین حالتی، بازیکن بدون توپ، مدل دنیای خود را به بازیکن صاحب توپ هم تیمی خود اعلام می کند و وی را در تعیین حالت و به تبع آن، تعیین عمل مناسب در آن حالت، یاری می دهد.

تغییر عمده دیگر در سناریوی شبیه سازیهای این قسمت، در نحوه قرار گیری بازیکن بدون توپ در زمانهایی است که بازیکن هم تیمی خود را در فاصله نزدیکتری به توپ می بیند. در چنین وضعیتی، بازیکن بدون توپ بر اساس همان تکنیک "بهترین گوشه در مربع حالت"، به مرکز مربع بلافاصله ای (از مربع حالت در بردارنده خود) می رود که آن مربع را جهت دریافت توپ (در صورت لزوم) از بازیکن هم تیمی صاحب توپ، مناسب تشخیص می دهد.

بدین نحو، در شبیه سازیهای این فصل، بازیکن بدون توپ نیز می تواند دارای حالات و اعمال خاص (حرکت به سمت یکی از ۸ جهت اطراف خود بجهت دریافت توپ) باشد و در واقع "حرکت بدون توپ" انجام دهد. در واقع با افزودن

به تعداد حالات و اعمال ممکن برای هر عامل نسبت به قسمت قبل، حرکات بازیکنان در این قسمت بسیار واقعی تر از قسمت قبل شده است. ضمن آنکه باید توجه داشت که عمل مناسب انتخابی در هر حالت در این قسمت، الزاما به معنی پاس ویا شوت نمی باشد و در بعضی موارد، معنی دریبل و در بعضی موارد نیز (در صورتی که بازیکن صاحب توپ، تنها در صورت داشتن عملی با ضریب اطمینان مثبت، آن را انجام دهد) معنی نگهداشتن موقت توپ بجهت یافتن روزنه ای برای انجام عمل مناسب را می دهد. نیز در روش ما، بازیکن خود را ملزم به حرکت رو به جلو (در همه حالات) نمی داند و در صورت لزوم از مفاهیم اوت کردن توپ، پاس به عقب، و حتی باز کردن بازی و ارسال توپ به فضای خالی (با احتمال دریافت مناسب توپ برای بازیکن هم تیمی و یا خود در چند سیکل بعد) نیز استفاده می کند. در واقع، ما از "تصمیم گیری بر روی فضای محدود و محلی" برای هر عامل استفاده کرده ایم. شبیه سازیهای ما نشان می دهند که این مورد در چنین دامنه پیچیده ای، جواب مناسبی می دهد[۱۳]. لازم بذکر است که یادگیری ما در کلیه شبیه سازیهای انجام گرفته کاملا چندعامله و توزیع شده است و مانند روشهای بکار رفته در اغلب مراجع، بازیکن صاحب توپ، قصد خود را از ارسال توپ به بازیکن بدون توپ اعلام نمی دارد (و همین طور بالعکس). بدین ترتیب، هر عامل در انتخاب عمل بهینه خود کاملا خودمختار است و هریک در روند انجام بازی و بدون استفاده از امکان ارتبا ، سعی می کنند که همکاری با یکدیگر را فرا بگیرند.در این مرحله، بد نیست به بازیهای انجام گرفته بین تیم نگاشت ثابت با تیم بدون یادگیری توجه شود.

در تیم نگاشت ثابت، بازیکنان در هر گوشه حالت که قرار داشته باشند، به سمت مرکز مربع نظیر همان گوشه، توپ خود را ارسال می کنند و در واقع، عمل آنها در هر لحظه وابسته به حالتی است که در آن لحظه در آن قرار دارند.مجموعا ۵۰ بازی بین این دو تیم برگزار شد که از مجموع این ۵۰ بازی، ۴ بازی مساوی، ۲۷ برد برای تیم نگاشت ثابت و ۱۹ برد برای تیم بدون یادگیری بوده است[۱۳]. بدین ترتیب مشخص است که تیم نگاشت ثابت خیلی بهتر از تیم بدون یادگیری عمل نمی کند و لهذا حالات و اعمال بهینه در آن حالات در روش ما ربطی به هم ندارند.

در شبیه سازی دیگری، تیم یادگیر Lrp با استفاده از تکنیک "بهترین گوشه در مربع حالت" را در چند بازی آموزشی در مقابل تیم $L_{2n,2}$ ۲ نفره که از عمومی سازی ساده حالات محیطی بخش قبل استفاده می کردند، قرار دادیم.

مجموعا ۳۰ بازی بین تیم Lrp با تیم $L_{2n,2}$ برگزار کردیم که نتیجه ۶ بازی اول آنها در زیر آمده است. در مجموع این ۳۰ بازی، ۲۸ بار تیم Lrp برنده بازی بود و یک بار نیز بازی به تساوی کشیده شد[۱۳]. به همین صورت در سری بعدی شبیه سازیها، ۳۰ بازی بین تیم Lrp با تیم "نگاشت ثابت" برگزار کردیم که در مجموع این ۳۰ بازی، ۲۴ بار تیم Lrp برنده بازی بود و یک بار نیز بازی به تساوی کشیده شد[۱۳].بدین ترتیب بار دیگر نتایج به ما نشان می دهند که این ایده که تعیین حالات ما با استفاده از این تکنیک، تعیین اعمال مناسب نظیر را نیز حاصل شود، اشتباه است.

سری شبیه سازیهای انجام شده برای تیمهای ۵ نفره و ۱۱ نفره و یادگیری تیمهای اتوماتای یادگیری که از این روش برای تعیین حالت عاملهای خود استفاده می کنند بر تیمهای بدون یادگیری و حتی دیگر تیمهای یادگیر[۱۲] نشان دهنده قابلیت تعمیم این روش به یک تیم کامل از عاملها می باشد.

۵. نتیجه گیری

بعنوان یک نتیجه کلی از شبیه سازیهای انجام شده در [۱۳] می توان گفت که روش بهترین گوشه در مربع حالت، عمومی سازی مناسبی از حالات محیط و نگاشت حالات یکسان به یک حالت واحد انجام می دهد. این مورد با توجه به غلبه تیمهای نمونه یادگیری که از این روش برای تعیین حالت عاملها استفاده می کنند بر تیمهای یادگیری که از دیگر روشها استفاده می نمایند، کاملا مشهود است.

شاید شاخص ترین مورد در این تکنیک، نگاشت فضای حالات پیوسته و بسیار وسیع به فضای حالاتی گسسته و محدود (بامشبک ساختن محیط دامنه عامل) برای عامل می باشد. ضمن آن که به شکل ظریفی نیز سعی شده است که مواردی چون فاصله و زاویه که مفاهیمی پیوسته هستند در تعیین حالت عامل (نسبت به اطراف خود و در محیط خود) در این تکنیک نقش داشته باشند.

نتایج شبیه سازیهای ما نشان می دهند که تکنیک بهترین گوشه در مربع حالت که باعث تصمیم گیری عامل بر روی فضای محدود و "محلی" خود در جهت همکاری با دیگر عاملها برای برآوردن هدف "سرتاسری" تیم است، بسیار کار "همکاری محلی بجهت دستیابی به هدف تیمی" سرتاسری می دهد و دامنه های پیچیده و همکاری گرای چند عامله بهره برد.

مراجع

- Weiss G., Multiagent Systems: A Modern Approach to istributed Artificial Intelligence, The MIT Press, London,)99.
- Stone P., Layered Learning in Multi_Agent Systems, PhD esis, School of Computer Science, Carnegie Mellon iiversity, December 1998.

Noda I., Team GAMMA: Agent programming on gaea, In :ano H., editor, RoboCup-97: Robot Soccer World Cup I, ges 500-507, Springer Verleg, Berlin, 1998.
RoboCup web page, 1997, At URL http://www.robocup.org.

Kitano H., editor, RoboCup-97: Robot Soccer World Cup I, ringer Verlag, Berlin, 1998.

Andre D., Corten E., Dorer K., Gugenberger P., Joldos M., mmenje J., Navaratil P. A., Noda I., Riley P., Stone P., :ahashi R., and Yeap T., Soccer server manual, version 4.0, hnical Report RoboCup 1998-001, RoboCup, 1998.

Narendra K.S. and Thathachar M.A.L., Learning Automata: Introduction, Prentice Hall, Inc., 1989.

Mars P., Chen J.R. and Nambir R., Learning Algorithms: ory and Applications in Signal Processing, Control and mmunications, CRC Press, Inc., pp. 5-24, 1996.

9- Lakshmivarahan S., Learning Algorithms: Theory and Applications, New York, Springer Verlag, 1981.

10- Meybodi M.R. and Lakshmivarahan S., ε — Optimality of a General Class of Absorbing Barrier Learning Algorithms, Information Sciences, Vol. 28, pp. 1-20, 1982.

11- Meybodi M.R. and Lakshmivarahan S., On a Class of Learning Algorithms which have a Symmetric Behavior under Success and Failure, Springer Verlag Lecture Notes in Statistics, pp. 145-155, 1984.

۱۲ – محمد رضا خجسته و محمد رضا میبدی/ اتوماتای یادگیر بعنوان مدلی برای همکاری در یک تیم از عاملها/ مرکز تحقیقات انفورماتیک/ آزمایشگاه محاسبات نرم/ دانشکده مهندسی کامپیوتر/ دانشگاه صنعتی امیرکبیر/ بهار ۱۳۸۱.

۱۳ - محمد رضا خجسته/ همکاری در سیستمهای چند عامله با استفاده از اتوماتای یادگیر/ پایان نامه کارشناسی ارشد/ دانشکده مهندسی کامپیوتر/ دانشگاه صنعتی امیرکبیر/ بهار ۱۳۸۱.

---

۱ behavior =

۲ cognitive –

۳ coordinate –

٤ opponents –

٥ Cooperative Team –

٦ Robocup –

۷ the robocup soccer server –

۸ real time –

۹ autonomous –

۱۰ noisy –

۱۱ collaborative –

۱۲ adversarial –

۱۳ hidden state –

۱٤ noisy –

۱٥ cycles –

۱٦ perception –

۱۷ action –

۱۸ unknown –

۱۹ realistic –

۲۰ Planned –

۲۱ existing –

۲۲ stamina –

۲۳ discrete –

۲٤ simulator cycle –

۲٥ marker –

۲٦ aural –

۲۷ visual –

۲۸ physical –

۲۹ crowded –

دانشگاه فردوسی مشهد

انجمن کامپیوتر ایران
Computer Society of Iran



8th Annual CSI Computer Conference

مجموعه مقالات

هشتمین کنفرانس بین المللی سالانه انجمن کامپیوتر ایران

گروه کامپیوتر - دانشکده مهندسی
دانشگاه فردوسی مشهد
۶ الی ۸ اسفند ۱۳۸۱

