

بسم الله الرحمن الرحيم



دانشگاه صنعتی اصفهان

دانشکده مهندسی برق و کامپیوتر

بهبود کیفیت و سرعت یادگیری در سیستم‌های چندعامله با استفاده از

ماتریس ارجاع و انتگرال فازی

پایان‌نامه کارشناسی ارشد مهندسی کامپیوتر – هوش مصنوعی و رباتیک

داریوش حسن‌پورآده

استاد راهنما

دکتر مازیار پالهنګ

۱۳۹۵

چهار



دانشگاه صنعتی اصفهان

دانشکده مهندسی برق و کامپیوتر

پایان نامه کارشناسی ارشد رشته مهندسی کامپیوتر – هوش مصنوعی و رباتیک آقای

داریوش حسن پور آده

تحت عنوان

بهبود کیفیت و سرعت یادگیری در سیستم‌های چندعامله با استفاده از

ماتریس ارجاع و انتگرال فازی

در تاریخ ... توسط کمیته تخصصی زیر مورد بررسی و تصویب نهایی قرار گرفت:

دکتر مازیار پالهننگ

۱- استاد راهنمای پایان نامه

دکتر ...

۳- استاد داور (اختیاری)

دکتر ...

۴- استاد داور (اختیاری)

دکتر محمد رضا تابان

سرپرست تحصیلات تکمیلی دانشکده

تشکر و قدردانی

پروردگار منّان را سپاسگزارم

کلیه حقوق مادی مترتب بر نتایج مطالعات،
ابتکارات و نوآوری‌های ناشی از تحقیق
موضوع این پایان‌نامه متعلق به دانشگاه
صنعتی اصفهان است.

دلشنکی‌های آدمی را باد ترانه‌ای می‌خواند

رویاهایش را آسمان پر ستاره نادیده می‌گیرد

و هر دانه‌ی برفی به اشکی نریخته می‌ماند.

سکوت سرشار از سخنان ناگفته است؛

از حرکات ناکرده،

اعتراف به عشق‌های نهان،

و شگفتی‌های به زبان نیامده،

دراین سکوت حقیقت ما نهفته است؛

حقیقت تو و من.

برای تو و خویش

چشمانی آرزو می‌کنم،

که چراغ‌ها و نشانه‌ها را در ظلمات مان ببیند.

گوشی،

که صداها و شناسه‌ها را در بیهوشی مان بشنود.

برای تو و خویش،

روحي،

که این همه را در خود گیرد و بپذیرد.

و زبانی

که در صداقت خود ما را از خاموشی خویش بیرون کشد،

و بگذارد از آن چیزها که در بندمان کشیده است، سخن بگوییم.

پنجه درافکنده‌ایم با دست‌هایمان

به جای رها شدن

سنگین سنگین بر دوش می‌کشیم

بار دیگران را

به جای همراهی کردن‌شان!

عشق، ما نیازمند دهان، است نه تصاحب

در راه خویش ایثار باید نه انجام وظیفه...

بی‌اعتمادی دری است

خودستایی، چفت و بست غرور است

و تهی دستی، دیوار است و لولا است

زندانی را که در آن محبوس رأی خویش ایم

دلتنگی مان را برای آزادی و دلخواه دیگران بودن

از رخنه‌هایش تنفس می‌کنیم...

- مارگوت بیگل

فهرست مطالب

| صفحه | عنوان |
|--------|---|
| یازده | فهرست مطالب |
| دوازده | فهرست تصاویر |
| ۱ | چکیده |
| ۲ | فصل اول : روش پیشنهادی |
| ۲ | ۱-۱ مقدمه |
| ۳ | ۲-۱ معیار خبرگی - ماتریس ارجاع و خاطره |
| ۶ | ۳-۱ یادگیری مشارکتی Q با استفاده از ماتریس ارجاع و انتگرال فازی |
| ۶ | ۱-۳-۱ الگوریتم پیشنهادی |
| ۷ | ۲-۳-۱ چرخه یادگیری مستقل |
| ۷ | ۳-۳-۱ چرخه همکاری |
| ۷ | ۴-۱ علت کارکرد انتگرال فازی چوکت در انتقال دانش |
| ۷ | ۱-۴-۱ ویژگی‌های انتگرال فازی چوکت |
| ۷ | ۲-۴-۱ علت عملکرد بهتر مدل انتگرال فازی نسبت به مدل رایج مجموع وزنی؟ |
| ۷ | مراجع |
| ۹ | چکیده انگلیسی |

فهرست تصاویر

چکیده

واژه‌های کلیدی: ۱- سیستم‌های چندعامله، ۲- یادگیری مشارکتی، ۳- یادگیری تقویتی، ۴- دانش غیرافزایشی، ۵- انتگرال فازی.

فصل اول

روش پیشنهادی

۱-۱ مقدمه

در این فصل جزییات روش پیشنهادی به طور مفصل معرفی خواهد شد، روش ارائه شده در حالت کلی از دو قسمت تشکیل شده است؛ اولین و مهم‌ترین قسمت ارائه یک معیار خبرگی جدید به نام معیار خبرگی «ارجاع» که برای هر عامل در هر چرخه یادگیری محاسبه و در یک «ماتریس ارجاع» نگه‌داری می‌شود. دومین قسمت مربوط به ترکیب دانش‌های عامل‌ها هستند که با استفاده از یک مدل انتگرال فازی، صورت می‌گیرد. همانطور که در فصل بعدی نیز نشان داده خواهد شد استفاده از مدل انتگرال فازی به دلیل خواصی مهمی که این مدل دارد باعث می‌شود سرعت و کیفیت یادگیری به طرز چشم‌گیری افزایش یابد. در این فصل ابتدا به معرفی معیار «ارجاع» و دلیل استفاده از آن می‌پردازیم سپس یادگیری مشارکتی چندعامله با استفاده از ماتریس ارجاع و انتگرال فازی معرفی خواهد شد و در نهایت نشان داده خواهد شد که چرا استفاده از انتگرال فازی نتایج بهتری را نسبت به مدل‌های سنتی چون مدل مجموع وزنی^۱ را ارائه می‌دهد.

^۱ Weighted Sum

۱-۲ معیار خبرگی - ماتریس ارجاع و خاطره

در دنیای واقعی «خبرگی» تعاریف متعددی به خود گرفته است، در روانشناسی خبرگی به معنی عملکرد برتر عامل تلقی می‌شود. در جامعه شناسی خبره به فردی گفتی برچسب خبرگی توسط یک گروهی به فرد زده شده است و آن گروه به توانایی که آن فرد در اختیار دارد علاقه‌مند^۱ است. در فلسفه خبره به فردی گفته می‌شود که دانشی که فرد تازه‌کار در اختیار ندارد را دارا می‌باشد [۱]. اگر تعاریف مختلف «خبرگی» را بررسی کنیم می‌بینیم که همه‌ی تعاریف در واقع تعبیری از میزان کیفیت عملکرد عامل نسبت به دیگر عامل‌ها می‌باشد. این تعبیر کلی از «خبرگی» انگیزه‌ای شد که درصدد معرفی معیاری برآیم که در حالت کلی بتوان به کلیه‌ی تعاریف «خبرگی» قابل تعمیم باشد.

تئوری ۱-۱. فرض می‌کنیم عامل A در محیط E در پی رسیدن به یک مجموعه اهداف $G \subseteq \{g_1, g_2, \dots, g_n\}$ می‌باشد. میزان خبرگی عامل رابطه‌ی معکوسی با میزان تلاش عامل برای رسیدن به اهداف تعریف شده خود دارد.

طبق آنچه که در تئوری بالا آورده شده است از بین چند عاملی که در یک محیط و یک مجموعه از اهداف فعالیت می‌کنند، عاملی خبره‌تر است که تلاش کمتری برای رسیدن به آن مجموعه اهداف می‌کند. شاید این مساله در نگاه اول نامتعارف به ذهن برسد ولی در فعالیت‌های روزمره ما انسان‌ها نیز به کرات شاهد این امر می‌باشیم. به عنوان مثال رانندگی دو فرد مبتدی و حرفه‌ای را در نظر بگیریم؛ فرد مبتدی هنگام رانندگی تمام حواس خود را معطوف به رانندگی می‌کند تلاش بسیار زیادی برای کنترل نسبت میزان کلاچ و گاز می‌کند و هنگام رانندگی به طور طبیعی رانندگی نمی‌کند و ... ولی فرد خبره کلیه موارد ذکر شده را بطور خودکار و طبیعی انجام می‌دهد بطوری که انگار رانندگی مانند دیگر رفتارهای طبیعی وی چون نفس کشیدن می‌باشد، که بصورت خودکار صورت می‌پذیرد. از این گونه مثال‌ها از کاربرد تئوری ۱-۱ در زندگی روزمره ما زیاد می‌توان یافت.

توجه شود که در تئوری ۱-۱ عبارت «میزان تلاش» عامل می‌تواند در کاربردهای مختلف تعبیر مختلفی به خود بگیرد، مثلاً در مثال راننده‌ی مبتدی و خبره میزان نسبت مسافت طی شده بر زمان رانندگی را می‌توان به عنوان «میزان تلاش» عامل در نظر گرفت که در شرایط یکسان راننده‌ی خبره‌تر به طور نسبی در زمان کوتاه‌تری یک مسافت مشخصی را طی خواهد کرد (در رد کردن پیچ و خم‌های ترافیک و مدت زمان ترمز و ... زمان کمتری را تلف می‌کند). یا به عنوان مثال دیگر، دانشجوی قوی و دانشجوی ضعیف را مورد بررسی قرار دهیم، دانشجویی خبره هست که زمان کمتری را صرف حل صحیح یک مساله خاص کند (با فرض اینکه دانشجویها حتماً باید مساله را حل کنند). همانطور که دیدیم کمیت «میزان تلاش» عامل برای مسائل مختلف معیار متفاوتی را دربر می‌گیرد ولی همگی از همان اصل معرفی شده در تئوری ۱-۱ تبعیت می‌کنند.

¹Interested

در یادگیری مشارکتی با استفاده از تئوری ۱-۱ می‌توان با تعریف ۱-۱ یک معیار خبرگی جدید را معرفی کرد که مبنی و پایه‌ی دستاوردهای این پژوهش می‌باشد.

تعریف ۱-۱ (معیار خبرگی «میزان ارجاع»). فرض می‌کنیم مجموعه‌ای از عامل‌ها $\mathbb{A} = \{A_1, A_2, \dots, A_m\}$ در محیط \mathcal{E} در پی رسیدن به یک مجموعه اهداف $\mathcal{G} \subseteq \{g_1, g_2, \dots, g_n\}$ می‌باشند. اگر ما به طور مجازی و دلخواه محیط \mathcal{E} را به k ناحیه مانند e_i افراز کنیم بطوری که $\mathcal{E} = \{\cup_{i=1}^k e_i \mid \forall i, j \in \{1, 2, \dots, k\} \wedge i \neq j : e_i \cap e_j = \emptyset\}$ طبق تئوری ۱-۱ در هر ناحیه i ام عاملی خبره‌تر است که میزان حضور آن عامل در آن ناحیه کمتر از دیگران است.

در تشریح آنچه که در تعریف ۱-۱ آمده است می‌توان گفت که در سیستم‌های چندعاملی که همگی عوامل در یک محیط به صورت مستقل در حال فعالیت هستند؛ محیط را به چند ناحیه دلخواه افراز می‌کنیم که اجتماع نواحی باهم کل محیط \mathcal{E} را تشکیل دهند و هیچ دو ناحیه‌ای اشتراکی باهم نداشته باشند [۲]. در این چنین افرازی از محیط، در هر ناحیه عاملی که نسبت به بقیه خبره‌تر است، نسبت به بقیه عوامل در همان ناحیه میزان تمایل حضور کمتری را از خود نشان می‌دهند. به عبارت دیگر عاملی که خبره‌تر است تمایل دارد کوتاه‌ترین مسیر رسیدن به اهداف خود را طی کند که نهایتاً منجر خواهد شد که میزان حضور عامل در هریک از نواحی محیط کمینه شود.

آنچه که در تئوری ۱-۱ در مورد «میزان تلاش» عامل آمده است در تعریف ۱-۱ در به صورت «میزان حضور عامل در هر ناحیه» تعریف شده است. بطوری که طبق تئوری مطرح شده میزان خبرگی عامل در هر ناحیه رابطه‌ی معکوسی با میزان حضور عامل در همان ناحیه را دارد. زیرا اگر عامل نسبت به محیط خود شناخت کامل‌تری داشته در هنگام تلاش برای رسیدن به اهداف خود به علت شناخت خوبی که از محیط دارد کمتر در محیط پرسه می‌زند (کمتر تلاش می‌کند) و با تعداد گام کمتری به سمت اهداف خود حرکت می‌کند - در واقع مسیر بهتری/کوتاه‌تری برای رسیدن به هدف را می‌شناسد. این موضوع در نهایت منجر می‌شود که عاملی که در هر ناحیه خبره‌تر است در همان ناحیه میزان پرسه زدن (حضور/تلاش) کمتری نسبت به دیگر عامل‌ها که از خبرگی نسبی کمتری برخوردار است را داشته باشد.

تا به اینجا گفته شد که عاملی که از خبرگی بیشتری برخوردار است لزوماً کمتر در محیط پرسه می‌زند و با طی کردن مسیر کوتاه‌تر به سمت اهداف خود، تلاش کمتری می‌کند ولی چند سوال در اینجا مطرح می‌شود که برای حل مساله نیازمند پاسخ به آن‌ها هستیم.

۱. میزان حضور عامل را در نواحی مختلف، که محیط از d -بعد تشکیل شده است چگونه مدل شود؟
۲. اگر عاملی که در هر چرخه یادگیری به یکی از نواحی کلا وارد نشد و میزان پرسه زدن عامل در آن ناحیه صفر شود؛ آیا این مقدار کمینه پرسه زدن، نشان دهنده‌ی خبرگی عامل در آن ناحیه است؟

۳. چگونه در معیار خبرگی ارائه شده باید مساله عدم حضور عامل در یکی از نواحی را مدل کرد، بگونه‌ای که اثر سوئی بر تجربه‌ی دیگر عامل‌ها در آن نواحی، در هنگام ترکیب دانش عامل‌ها نداشته باشد؟

پاسخ به این سوالات برای حل مساله با استفاده از معیار خبرگی پیشنهادی (تعریف ۱-۱) ضروری است. ما به ازای کلیه‌ی نواحی یک ماتریسی به نام «ماتریس ارجاع» (یا به اختصار REFMAT^۱) در نظر میگیریم که در ابتدا صفر مقداردهی شده‌اند و هر دفعه که عامل از موقعیتی به موقعیت دیگر می‌رود مقدار آن ناحیه‌ای که موقعیت جدید در آن واقع است را یک واحد افزایش می‌دهیم بدین وسیله میزان حضور عامل در نواحی مختلف را می‌شماریم. همانطور که در قسمت آزمایشات این پایان‌نامه نشان داده شده است در صورتی که از تابع انتخاب عمل بولترمن استفاده کنیم میزان کوچک یا درشت بودن این نواحی در کیفیت نتیجه تاثیرگذار نیست. یعنی عملاً چه ما در حالت کلی، کل محیط را به عنوان یک ناحیه در نظر بگیریم و میزان حضور عامل در این ناحیه را بشماریم (که معادل می‌شود با تعداد گام‌های عامل در طی رسیدن به هدف) یا در حالت جزئی به ازای هر موقعیت موجود را یک ناحیه در نظر بگیریم (که معادل می‌شود با تعداد ملاقات هر یکی از موقعیت‌ها توسط عامل) به یک نتیجه می‌رسیم.

به همین دلیل در پاسخ به سوال دوم، اگر تعداد نواحی زیاد باشد (مثلاً هر موقعیت یک ناحیه باشد - حداکثر تعداد نواحی) ممکن است عامل در طی رسیدن به هدف برخی از نواحی را کلاً ملاقات نکند و مقدار ارجاع به آن نواحی صفر شود و از طرفی طبق تعریف ۱-۱ عاملی که تعداد حضور کمتری در نواحی مختلف داشته باشد از خبرگی بیشتری در آن نواحی برخوردار است و در این شرایط که مقدار ارجاع عامل به ناحیه‌ای صفر باشد را نمی‌توان به خبرگی عامل در آن ناحیه نسبت داد زیرا که آن عامل در کل، آن ناحیه را ملاقات نکرده است که بخواهد تجربه‌ای را در تعامل با آن ناحیه کسب کند تا بتواند خبرگی خود را در آن ناحیه افزایش دهد. برای حل این مشکل و پاسخ به سوال سوم، ماتریسی جدیدی به نام ماتریس خاطره (یا به اختصار RCMAT^۲) را معرفی می‌کنیم. این ماتریس وظیفه‌ی نگهداری آخرین ارجاعات غیر صفر عامل را به هر کدام از نواحی تعریف شده را دارد و در زمان‌هایی که مقدار یک ناحیه در ماتریس REFMAT صفر باشد مقدار آن ناحیه از ماتریس RCMAT بروز رسانی می‌شود که میزان پرسه زدن عامل در آن ناحیه در آخرین باری عامل آن ناحیه را ملاقات کرده است را می‌دهد؛ در صورتی که مقدار پرسه زدن یک ناحیه در ماتریس REFMAT مقداری غیر صفر باشد مقدار ماتریس RCMAT با مقدار کنونی REFMAT آن ناحیه بروز رسانی می‌شود.

دلیل استفاده از ماتریس RCMAT این است که در یادگیری تقویتی عامل زمانی می‌توان دانش (سیاست/خبرگی)

^۱Reference Matrix

^۲Recall Matrix

خود را نسبت به نحوه‌ی عمل در یک موقعیت بهبود ببخشد که آن موقعیت را ملاقات کند. حال اگر عامل موقعیتی را ملاقات نکند دانش وی در آن موقعیت ثابت خواهد ماند به همین دلیل اگر عامل ناحیه‌ای را ملاقات نکند و مقدار REFMAT آن ناحیه صفر باشد می‌دانیم که دانش (خبرگی) عامل در آن ناحیه در این چرخه‌ی یادگیری ثابت مانده است و در صورتی که دوباره در آن ناحیه قرار می‌گرفت، حدودا به همان میزان آخرین ملاقات در آن محیط پرسه خواهد زد. به عبارت دیگر در یک چرخه یادگیری اگر هر ناحیه ملاقات نشده، مورد ملاقات واقع می‌شد، تقریباً به میزان آخرین تعداد ارجاع شده برای آن نواحی، مورد ارجاع واقع می‌شد.

۱-۳ یادگیری مشارکتی Q با استفاده از ماتریس ارجاع و انتگرال فازی

آنچه که تا به اکنون در مورد روش پیشنهادی این پژوهش آورده شده، معرفی یک معیار خبرگی که در برعکس بسیاری از معیارهای خبرگی که تا به کنون معرفی شده است [۳-۵] در تمامی موقعیت‌های دنیای واقعی به وفور مشاهده می‌شود و آن ارائه این تئوری است عامل خبره‌تر برای رسیدن به یک مجموعه از اهداف تلاش نسبی کمتری نسبت به دیگر عامل‌ها با خبرگی کمتر در شرایط یکسان می‌کند. حال که معیاری برای میزان خبرگی عامل‌ها در اختیار داریم چالش بعدی برای بهبود کیفیت و سرعت یادگیری مشارکتی ارائه‌ی روشی برای ترکیب دانش‌های عامل‌ها از محیط (جداول Q آن‌ها) با استفاده از معیار ارائه شده می‌باشد. روش ترکیب باید بگونه‌ای باشد که کیفیت و سرعت یادگیری مشارکتی عامل‌ها را در طی زمان نسبت زمانی که عامل‌ها بدون مشارکت یاد می‌گیرند بهتر کند. همچنین کیفیت و سرعت یادگیری همبستگی مستقیمی داشته باشند با تعداد عامل‌هایی که در حال اشتراک گذاری هستند؛ به عبارت دیگر در صورت افزایش تعداد عامل‌هایی که دانش‌های خود را به اشتراک می‌گذارند مدل ترکیب کننده‌ی دانش‌های آن عامل‌ها باید بتواند دانش بهتری تولید کند که نهایتاً منجر به بهتر شدن کیفیت و سرعت کلی یادگیری عامل‌ها شود.

در این پژوهش ما انتگرال فازی را به عنوان مدل ترکیب کننده‌ی دانش‌های عامل‌ها پیشنهاد می‌دهیم. دلیل انتخاب این مدل ویژگی‌های منحصر به فردی است که این مدل کننده در اختیار دارد که مدل را کاملاً مناسب برای ترکیب دانش عامل‌ها می‌کند؛ که در بخش‌های آتی این فصل بطور مفصل توضیح می‌دهیم که این ویژگی‌ها چه هستند و چرا این ویژگی‌ها برای ترکیب دانش عامل‌ها مناسب هستند.

۱-۳-۱ الگوریتم پیشنهادی

نپن

۱-۳-۲ چرخه یادگیری مستقل

۱-۳-۳ چرخه همکاری

۱-۴-۴ علت کارکرد انتگرال فازی چوکت در انتقال دانش

۱-۴-۱ ویژگی‌های انتگرال فازی چوکت

هیچ وقت از حدود تجاوز نمی‌کند

مناسب برای اندازه‌گیری‌های غیرافزایشی

۱-۴-۲ علت عملکرد بهتر مدل انتگرال فازی نسبت به مدل رایج مجموع وزنی؟

مراجع

- [1] “Expert - wikipedia.” <https://en.wikipedia.org/wiki/Expert>. (Accessed on 11/12/2016).
- [2] E. Schechter, *Handbook of Analysis and its Foundations*, ch. 1, p. 16. Academic Press, 1996.
- [3] M. N. Ahmadabadi, M. Asadpur, S. H. Khodanbakhsh, and E. Nakano, “Expertness measuring in cooperative learning,” in *Intelligent Robots and Systems, 2000.(IROS 2000). Proceedings. 2000 IEEE/RSJ International Conference on*, vol. 3, pp. 2261–2267, IEEE, 2000.
- [4] E. Pakizeh, M. Palhang, and M. M. Pedram, “Multi-criteria expertness based cooperative q-learning,” *Applied intelligence*, vol. 39, no. 1, pp. 28–40, 2013.
- [5] M. ali mirzaei badizi, “Speed-up cooperative learning in multi-agent systems using shortest experimented path,” Master’s thesis, Department of Electrical and Computer Engineering, Isfahan University of Technology, Isfahan University of Technology, Isfahan 84156-83111, Iran, 3 2015.

Improvements in speed and quality of learning in multi-agent systems using the reference matrix and fuzzy integral

Dariush Hasanpour Adeh

d.hasanpoor@ec.iut.ac.ir

[DATE]

Department of Electrical and Computer Engineering
Isfahan University of Technology, Isfahan 84156-83111, Iran

Degree: M.Sc.

Language: Farsi

Supervisor: Assoc. Prof. Maziar Palhang (palhang@cc.iut.ac.ir)

Abstract

Key Words:

Multi-agent Systems, Cooperative Learning, Reinforcement Learning, Non-additive Knowledges, Fuzzy Integral



Isfahan University of Technology

Department of Electrical and Computer Engineering

Improvements in speed and quality of learning in multi-agent systems using the reference matrix and fuzzy integral

A Thesis

Submitted in partial fulfillment of the requirements
for the degree of Master of Science

by

Dariush Hasanpour Adeh

Evaluated and Approved by the Thesis Committee, on ...

1. Maziar Palhang, Assoc. Prof. (Supervisor)
2. ..., Prof. (Examiner)
3. ..., Prof. (Examiner)

Mohamad Reza Taban, Department Graduate Coordinator

