Downloaded from adst.ir at 11:32 +0430 on Saturday August 12th 2017

مجله علمی پژوهشی «علوم و فناوری‌های پدافند نوین»
سال هشتم، شماره ۲، تابستان ۱۳۹۶؛ ص ۱۵۹-۱۴۷

طراحی سرویس طرح‌ریزی مشارکتی برخط مبتنی بر فرایند مارکوف در حوزه فرماندهی و کنترل

مهدی نقیان فشارکی*[۱]، سعیده ساداتی[۲]، امیرحسین مؤمنی ازندریانی[۳]، سید مصطفی حسینی[۴]
۱- استادیار، ۲- کارشناس ارشد، ۳- دانشجوی دکتری، دانشگاه صنعتی مالک اشتر، ۴- دانشجوی دکتری، دانشگاه صنعتی امیرکبیر
(دریافت: ۹۳/۰۱/۲۵، پذیرش: ۹۵/۰۸/۳۰)

چکیده

رویکردهای طرح‌ریزی موجود اگر چه می‌توانند در شرایط عدم قطعیت و به صورت غیر متمرکز طرح‌ریزی نمایند، اما اکثر آن‌ها در مواقعی که شرایط پیچیده سناریوهای فرماندهی و کنترل همچون نیاز به طرح‌ریزی دقیق، تصمیم‌گیری بلادرنگ، ارتباطات محدود بین عامل‌ها و پویایی محیط حاکم است، بازده خوبی نداشته و گاهی با شکست مواجه می‌شوند. از بین رویکردهای موجود، مدل‌های توسعه یافته DEC-POMD مانند MAOP-COMM، برای این شرایط مناسب هستند. می‌توان با بهبود مدل MAOP-COMM، طرح‌ریزی دقیق‌تری برای عامل‌ها انجام داد. در این مقاله با ارتقاء اکتشاف تابع ارزش و به‌کارگیری" پیش‌بینی دومرحله‌ای" در استراتژی یافتن خط مشی بهینه و اخذ تصمیم صحیح، الگوریتم اخیر بهبود داده شد. در مرحله بعد رویکرد بهبود یافته را در قالب یک وب سرویس پیاده‌سازی شد تا برای هر سناریوی دلخواه با شرایط محدود شده، سناریو طراحی شده را از معمار سناریو دریافت کند و با فرض پویا بودن محیط، طرح‌ریزی لازم را به طور برخط انجام داده و دنباله اقدامات عامل‌ها را در هر لحظه تعیین نماید. نتایج به‌دست آمده، برتری الگوریتم بهبود یافته را نشان می‌دهد، به گونه‌ای که سرویس طراحی شده با الگوریتم بهبود یافته می‌تواند در شرایط عدم قطعیت به صورت غیر متمرکز و برخط با کارایی مضاعف و درصد بالایی از صحت تصمیم‌گیری برای عامل‌ها طرح‌ریزی نماید.

کلیدواژه‌ها: طرح‌ریزی مشارکتی، فرماندهی و کنترل، طرح‌ریزی برخط، معماری سرویس‌گرا، فرایند تصمیم مارکوف

Design Online Collaborative Planning Service Based on Markov Process in Command and Control Domain

M. Naghian Fesharaki*, S. Sadati, A. H. Momeni Azndarian, S. M. Hosseini
Malek Ashtar University of Technology
(Received: 14/04/2014; Accepted: 20/11/2016)

Abstract

Although existing Planning methods can plan under uncertainty and decentralize situation, most of them malfunction in some complicated conditions of command and control scenarios such as real time decision making, needing more accurate planning, bounded communication between agents, dynamic worlds and partially observable environments. Among suitable models for these situations, extended models of DEC-POMDPs such as MAOP-COMM can be considered for handing these conditions. It is possible to improve MAOP-COMM model to do planning for agents with double precision. In this paper, the algorithm of MAOP-COMM model has been improved by upgrading value function heuristic and using "two steps look ahead" in the strategy of finding best policy and making correct decision. In the next step, the improved method was implemented in a web service frame work. The web service gets the desired bounded scenario from the architect and performs online planning to generate plans and sequence of actions regarding to environment dynamicity. The obtained results, show preference of the improved algorithm, So that the designed service with improved algorithm can perform online planning for agents in a decentralized multi-agent system in uncertain condition with better performance and high percent of correct decision making.

Keywords: Collaborative Planning, Command and Control, Online Planning, Service Orient Architecture, Markov Decision Process.

*Corresponding Author E-mail: Mehfesharaki@mut.ac.ir

Archive of SID
www.SID.ir

Downloaded from adst.ir at 11:32 +0430 on Saturday August 12th 2017

۱. مقدمه

سناریوهای امروزی به دلیل پیچیده و پویا بودن صحنه‌ها با سناریوهای سنتی بسیار متفاوت شده و نیازمند فرماندهی و کنترل چابک و پویا هستند تا بتوانند به سرعت در برابر تغییرات صحنه نبرد واکنش نشان داده و تصمیم‌گیری نمایند. در محیط‌های فرماندهی و کنترل، طرح‌ریزی، از جمله مهم‌ترین اقدامات زمینه‌ای به شمار می‌رود. طرح‌ریزی، فرایند ایجاد شرایط لازم برای هماهنگ‌سازی اقدام‌ها و تلاش‌ها جهت دستیابی به هدفی واحد است [۱]. در محیط‌های چند عاملی هر عامل باید محدودیت ناشی از اقدامات عامل‌های دیگر را نیز در طرح‌ریزی در نظر بگیرد. از این رو فرایند طرح‌ریزی چند عاملی پیچیده‌تر از طرح‌ریزی تک عاملی است. طرح‌ریزی چند عاملی می‌تواند متمرکز یا غیر متمرکز باشد. منظور از طرح‌ریزی مشارکتی، طرح‌ریزی چند عاملی غیر متمرکز است. در طرح‌ریزی غیر متمرکز هر عامل به طور مستقل برای خودش تصمیم می‌گیرد که بر اساس اطلاعاتی که دریافت کرده یا مرتبط به او است، چه اقدامی انجام دهد یا چه طرحی پیاده‌سازی کند [۲]. طرح‌ریزی در دنیای چند عاملی را می‌توان به روش‌های مختلفی مدل کرد. از جمله این روش‌ها می‌توان عامل‌های شناختی BDI[1]، تئوری بازی‌ها، فرآیندهای تصمیم مارکوف[2] و مدل‌های توسعه یافته آن از جمله POMDP، DEC-POMDP[3] را نام برد.

محیط‌های فرماندهی و کنترل ویژگی‌ها و پارامترهای خاصی دارند، از جمله مهم‌ترین این پارامترها می‌توان به عدم قطعیت، پویایی، نقص اطلاعات، تنوع در وظایف عامل‌ها، بلادرنگ بودن و نیاز به طرح‌ریزی بدون مدیریت مرکزی اشاره کرد. بنابراین برای داشتن یک سامانه طرح‌ریزی جامع در محیط‌های فرماندهی و کنترل باید به دنبال رویکرد طرح‌ریزی مناسبی بود که بتواند علاوه بر پوشش دادن اکثر پارامترهای مذکور، ما را به تصمیم‌گیری‌های دقیق نیز برساند. از بین مدل‌های طرح‌ریزی موجود، رویکردهای توسعه یافته مدل DEC-POMDP مانند MAOP-COMM با پارامترهای یاد شده انطباق نسبی دارند.

از این رو در این مقاله روش فرآیندهای تصمیم‌گیری مارکوف نیمه رؤیت پذیر غیر متمرکز[4] اتخاذ می‌شود تا تیمی از عامل‌های همکار را که با یک محیط تصادفی نیمه رؤیت پذیر در ارتباط هستند، مدل کرد. از جمله جدیدترین توسعه مدل‌های DEC_POMDP که برخلاف سایر توسعه‌ها، از پارامتر برخط بودن نیز پشتیبانی می‌کند، می‌توان به مدل MAOP-COMM اشاره کرد . الگوریتم‌های بر خط، با دریافت تمام اطلاعات موجود، در هر لحظه، فقط یک گام طی می‌کنند [۳].

رویکرد MAOP-COMM علاوه بر اینکه طرح‌ریزی را به صورت برخط انجام می‌دهد، با محدود کردن ارتباطات عامل‌ها هماهنگی بین عامل‌ها را نیز تضمین می‌کند. اما این الگوریتم به دلیل ضعف تابع هیوریستیک و استفاده از پیش‌بینی یک مرحله‌ای بازده خوبی در اتخاذ خط مشی بهینه و انتخاب اقدامات برگزیده ندارد. در این مقاله سعی شده است تا با ارتقاء تابع هیوریستیک و افزایش گام‌های پیش‌بینی، بازده بسیار بهتری در اتخاذ خط مشی بهینه و انتخاب اقدامات برگزیده حاصل شود. در گام بعد رویکرد بهبود یافته در قالب یک وب سرویس پیاده‌سازی می‌شود تا برای هر سناریوی دلخواه با شرایط محدود شده، سناریو طراحی شده را از معمار سناریو دریافت کند و با فرض پویا بودن محیط، طرح‌ریزی لازم را به طور برخط انجام داده و دنباله اقدامات صحیح را برای عامل‌ها در هر لحظه ایجاد نماید.

در ادامه ساختار مقاله به صورت زیر آورده شده است: در بخش دوم ویژگی‌های بارز محیط‌های پیچیده و مدل‌های طرح‌ریزی مشارکتی را به اختصار بیان کرده، مدل‌ها را از حیث پارامترهای محیط‌های پیچیده فرماندهی و کنترل دسته‌بندی می‌شود. در بخش سوم پس از معرفی مدل MAOP-COMM به نحوه طرح‌ریزی این مدل اشاره می‌شود و بهبودهای پیشنهادی صورت گرفته بر این مدل را مطرح می‌شود. در بخش چهارم جزییات طراحی سرویس طرح‌ریزی مشارکتی بیان می‌شود. در بخش پنجم نیز نتایج آزمایش‌های انجام شده، نحوه به‌کارگیری سرویس طراحی شده و محدودیت‌های آن ذکر می‌شود.

۲. طرح‌ریزی مشارکتی در حوزه فرماندهی و کنترل

در این بخش ابتدا اهم ویژگی‌ها و شرایط محیط‌های فرماندهی و کنترل را بررسی کرده، به طور خلاصه مدل‌های موجود برای طرح‌ریزی مشارکتی را بیان می‌شود و سپس از بین مدل‌ها مدلی که به حوزه فرماندهی و کنترل نزدیک‌تر است معرفی می‌شود.

۲-۱. ویژگی‌های بارز محیط‌های فرماندهی و کنترل

اگر یک سامانه طرح ریز هوشمند بخواهد برای یک سناریو در محیط پیچیده فرماندهی کنترل، طرح‌ریزی کند و طرح حاصل مفید و معقول باشد، باید ویژگی‌های اساسی این حوزه را مورد توجه قرار داده و در مدل خود لحاظ نماید. در زیر اهم این ویژگی‌ها که با توجه به نیازمندی‌های حوزه فرماندهی و کنترل از منابع گوناگون استخراج شده‌اند، ذکر شده است. البته باید به این امر توجه نمود که الزاماً همه این ویژگی‌ها در همه حوزه‌ها مطرح نیستند و معمار دامنه بسته به نیاز خود با چند مشخصه یا همه این مشخصه‌ها مواجه است.

[1] Belief Desire Intention
[2] MDP
[3] Decentralize Partially Observable Markov Decision Processes
[4] DEC-POMDP

www.SID.ir

Downloaded from adst.ir at 11:32 +0430 on Saturday August 12th 2017

عدم قطعیت[1]: میدان دید محدود (نیمه رؤیت‌پذیری[2])، ترکیب مجموعه اطلاعات توأم با نویز و عدم قطعیت حسگرها علت اصلی عدم قطعیت است که لزوم کسب بهترین تخمین را در سناریوهای فرماندهی و کنترل موجب می‌شود [۲].

تصمیم‌گیری بلادرنگ[3]: کسب و پردازش اخبار و داده‌ها در صحنه‌ها زمان‌بر است. می‌توان گفت، فرماندهی و کنترل یک مسابقه در مقابل زمان است. بنابراین، از جمله مهم‌ترین نیاز مطلق در هر سامانه فرماندهی و کنترل، سرعت تصمیم‌گیری است [۴].

چند عاملی: در محیط‌های فرماندهی و کنترل به ندرت پیش می‌آید که یک عامل به تنهایی و مستقل ایفای نقش کند و همیشه چندین عامل در قالب یگان‌های عملیاتی، پست‌های فرماندهی، شبکه‌های همکار، گروه‌های تصمیم‌گیری و غیره به صورت مشارکتی با یکدیگر تعامل می‌کنند تا هماهنگ با هم به هدف مشترکی دست یابند [۵].

عدم تمرکز[4]: پیچیده شدن شرایط حوزه‌های فرماندهی و کنترل، سبب شده است تا عملیات‌ها گاهی در وضعیت حذف فرمانده مرکزی و به صورت توزیع شده قرار گیرند. به گونه‌ای که عامل‌های صحنه باید با مشارکت یکدیگر و با تصمیم‌گیری‌های مستقل برای دست‌یابی به هدفی واحد طرح‌ریزی کنند [۵].

تنوع عامل‌ها[5]: عامل‌های همکار همواره وظایف مشابه ندارند و ممکن است وظایف و اقدامات متنوعی داشته باشند و برخی اقداماتی که یک عامل می‌تواند انجام دهد مختص همان عامل باشد.

ارتباطات محدود: اطلاعات اغلب به وسیله پهنای باند محدود می‌شود و گاهی می‌تواند هزینه‌بر و غیر قابل اعتماد باشد. حتی وقتی ارتباط واقعاً موجود و ارزان است، پهنای باند محدود و عدم اطمینان اغلب منجر به تأخیر می‌شود. از این رو در محیط‌های پیچیده، ارتباطات محدود ارجحیت دارد [۳].

پویایی: در دامنه‌های واقعی حوزه فرماندهی و کنترل، محیط از پویایی بالایی برخوردار است. به این معنا که وضعیت دنیای پیرامون همواره در حال تغییر و احتمالاً غیر قابل پیش‌بینی است. از این رو بروز شدن مداوم همراه با محیط، چالش مهمی را در فرماندهی و کنترل باعث می‌شود [۵].

[1] Uncertainty
[2] Partially Observable
[3] Real Time Decision Making
[4] Decentralize
[5] Diversity

۲-۲. رویکردهای طرح‌ریزی مشارکتی

مدل‌های زیادی برای طرح‌ریزی در عدم قطعیت وجود دارد. اما همه آن‌ها از ویژگی‌های محیط‌های پیچیده پشتیبانی نمی‌کنند. هدف ما تمرکز بر مسائل طرح‌ریزی چند عاملی در شرایط عدم قطعیت[6] است که از مسائل تصمیم‌گیری غیر متمرکز[7] مشارکتی به شمار می‌روند. در ادامه مهم‌ترین رویکردهای طرح‌ریزی چند عاملی را مرور کرده، آن‌ها را از جهت پشتیبانی از پارامترهای ذکرشده در بخش قبل بررسی می‌شود.

رویکرد تئوری بازی‌ها: تئوری بازی مطالعه تصمیم‌گیری استراتژیک است. از جنبه دیگر تئوری بازی‌ها مطالعه مدل‌های ریاضی تضاد و هماهنگی بین تصمیم‌گیرنده‌های هوشمند است و برای وضعیت‌هایی که خروجی یک عامل علاوه بر خودش به انتخاب سایر عامل‌ها نیز وابسته است، طراحی شده است [۶]. مدل تئوری بازی‌ها را معمولاً به صورت درخت‌های بازی نشان می‌دهند. در این مدل عامل‌ها درباره هزینه تصمیم‌گیری یا ارتباط خود استدلال می‌کنند و می‌توانند مسائل تصمیم‌گیری دنباله‌ای را در محیط غیر قطعی مدل کنند. اما درخت بازی آن برای مدل کردن محیط‌های فرماندهی و کنترل بسیار بزرگ می‌شود. تئوری بازی‌ها اگر چه برای مدل کردن مسائل در عدم قطعیت مناسب است، اما سازوکاری برای برخورد با نیمه رؤیت پذیری ندارد. از این رو برای مسائل دنیای واقعی پویایی کافی نخواهد داشت مگر آنکه یک بازیکن اقدام تصادفی که حرکت شانس را تولید می‌کند به آن اضافه شود [۷]. به علاوه در این مدل فرض می‌شود که همه عامل‌ها مشاهدات کامل دارند و نیازی به ارتباط برای تبادل مشاهدات ندارند. از این رو، این مدل برای سناریوهای با میدان دید محدود مناسب نیست.

مدل BDI[8]: مدل نرم‌افزاری BDI (باور- مطلوب- تصمیم) یک مدل نرم‌افزاری است که برای برنامه‌نویسی عامل‌های هوشمند توسعه یافته است. به طور کلی، مدل BDI تلاشی برای حل مسئله‌ای است که بیشتر با طرح‌ها و طرح‌ریزی حل می‌شود تا با برنامه‌نویسی عامل‌های هوشمند، و از لحاظ ظاهری به وسیله پیاده‌سازی تصورات، خواسته‌ها و باورهای عامل مشخص می‌شود. این مدل مفاهیم را برای حل مسئله خاص در برنامه‌نویسی چند عاملی به‌کار می‌برد. اساساً این مدل سازوکاری برای جدا کردن عمل انتخاب یک طرح (از یک کتابخانه طرح یا فعالیت طرح‌ریز خارجی) از اجرای طرح‌های جاری ارائه می‌دهد. به منظور دست‌یابی به این جداسازی مدل نرم‌افزاری BDI جنبه‌های قاعده‌مند تئوری استدلال نظری انسان را پیاده‌سازی کرد [۸].

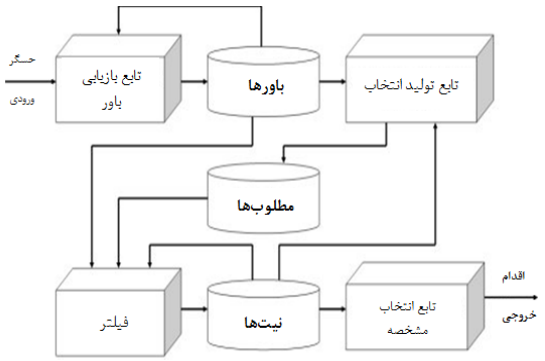
[6] Uncertainty
[7] Decentralize
[8] Belief Desire Intention

Arhive of SID

www.SID.ir

Downloaded from adst.ir at 11:32 +0430 on Saturday August 12th 2017

یک "عامل BDI" نوع خاصی از عامل نرم‌افزاری منطقی است که با نگرش‌های ذهنی خاصی ترکیب شده است که به اختصار باورها[1]، مطلوب‌ها[2] و تصمیم‌ها[3] خوانده می‌شوند. عامل‌های BDI قادرند تا توازنی بین زمان سپری شده برای تفکر درباره طرح‌ها (انتخاب اینکه چه کنند) و انجام طرح (اجرای آن) برقرار کنند. فعالیت بعد، تولید طرح در اولین مکان (طرح‌ریزی) است. در (شکل (۱)) نمودار مدل BDI نشان داده شده است.



شکل ۱. نمودار مفسر عامل BDI [۸]

ساختار عامل‌های BDI: اجزای ساختاری دلخواه برای یک سامانه BDI عبارت است از:

Beliefs: باورها نشان دهنده حالت اطلاعاتی یک عامل هستند، به بیان دیگر، باور عامل درباره جهان (شامل خودش و دیگر عامل‌ها) را بیان می‌کنند.

Belief Set: باورهایی است که در پایگاه داده ذخیره می‌شوند (گاهی پایگاه باور یا مجموعه باور خوانده می‌شود).

Desires: مطلوب‌ها حالت انگیزشی یک عامل را بیان می‌کنند. مطلوب‌ها اهداف یا وضعیت‌هایی را نشان می‌دهند که عامل‌ها می‌خواهند محقق شود یا به آن برسند.

Goals: هدف مطلوبی است که برای پیگیری فعال توسط یک عامل پذیرفته شده است. استفاده از اصطلاح هدف این محدودیت را در آینده به وجود می‌آورد که مجموعه مطلوب‌های فعال باید سازگار باشند.

Intentions: تصمیم‌ها نشان دهنده حالت مشورتی و تفکری یک عامل است –آنچه یک عامل باید انجام دهد. تمایلات مطلوب‌هایی هستند که عامل‌ها باید به صورت گسترده‌ای انجام دهند.

Events: رخدادها راه‌اندازهایی برای فعالیت‌های تراکنشی هستند. یک رخداد می‌تواند طرح‌ها را به‌روز کند، طرح‌ها را راه‌اندازی کند یا اهداف را اصلاح نماید. رخدادها می‌توانند به صورت خارجی رخ دهند و به وسیله حسگرها یا سامانه‌های مجتمع دریافت شوند. به علاوه عامل‌ها می‌توانند به صورت داخلی برای راه‌اندازی به‌روزرسانی‌های تفکیکی۴ یا اهداف اصلاح شده ایجاد شوند [۹].

مدل نرم‌افزاری BDI نمونه‌ای از ساختار استدلالی برای یک عامل منطقی منفرد است و در سامانه چند عاملی گسترده نگران کننده است. در زیر بخشی از نگرانی‌ها راجع به مدل نرم‌افزاری BDI با برجسته کردن محدودیت‌های شناخته شده ساختار بیان شده است:

- یادگیری: عامل‌های BDI از فقدان سازوکار مشخصی در ساختار برای یادگیری از رفتار پیشین و انطباق با وضعیت‌های آتی رنج می‌برند [۱۰ و ۱۱].
- سه وضعیت: نظریه پردازان تصمیم‌گیری و محققان طرح‌ریزی از ضرورت داشتن همه سه وضعیت می‌پرسند و محققان هوش مصنوعی توزیع شده می‌پرسند که آیا این سه وضعیت کافی است؟ [۹]
- ارتباط در منطق: منطق‌های چند وجهی که BDI را شامل می‌شوند، در عمل ارتباط کمی دارند [۹]. بنابراین مدل BDI به طور صریح سازوکاری برای تعامل با دیگر عامل‌ها جهت همگرایی در سامانه چند عاملی ندارد [۱۲].
- اهداف صریح: اکثر پیاده‌سازی‌های BDI نمایش واضحی از اهداف ندارند [۱۳].
- پیش‌بینی: این ساختار هیچ بررسی پیش‌بینی یا طرح‌ریزی پیش رو (در طراحی) انجام نمی‌دهد. این امر ممکن است نامطلوب باشد چرا که طرح‌های اتخاذ شده ممکن است منابع محدود را مصرف کنند، اقدامات ممکن است برگشت‌پذیر نباشند، اجرای یک وظیفه بیشتر از طرح‌ریزی پیش رو طول بکشد و در نهایت، اقدامات ممکن است منجر به عدم موفقیت شوند [۸].

رویکرد DCOP: رویکرد DCOP مسئله‌ای است که گروهی از عامل‌ها باید به صورت توزیع شده مقداری را برای مجموعه‌ای از متغیرها انتخاب کنند به طوری که هزینه یک مجموعه از محدودیت‌ها روی متغیرها بیشینه یا کمینه شود. این چارچوب برای توصیف یک مسئله (که به صورت محدودیت‌های شناخته شده‌ای توسط عامل‌ها در دست اجرا است) به‌کار می‌رود. محدودیت‌ها روی برخی متغیرها با دامنه‌های از پیش تعریف شده توصیف می‌شوند و باید توسط عامل به مقادیر مشابهی منتسب

[1] Belief
[2] Desire
[3] Intention

[4] Decoupled

Archive of SID

www.SID.ir

Downloaded from adst.ir at 11:32 +0430 on Saturday August 12th 2017

شوند. در مدل DCOP، سامانه چند عاملی به صورت یک مسئله بهینه‌سازی گسسته که بین عامل‌ها توزیع شده است، نمایش داده می‌شود [۱۴].

مدل DCOP برای محیط‌های پویا مناسب نیست. همچنین نمی‌توان این مدل را در محیط‌های نیمه رؤیت پذیر و غیر قطعی به‌کار برد [۱۵]. چرا که این مدل سازوکاری برای مدل کردن عدم قطعیت در خود ندارد. به علاوه انتساب مقادیر در محیط نیمه رؤیت پذیر با دشواری‌هایی به همراه خواهد بود.

مدل‌های طرح‌ریزی مبتنی بر فرآیند مارکوف: این مدل‌ها در شرایطی که چندین عامل با یکدیگر و احتمالاً روی کارهای مختلف کار می‌کنند تا به هدف مشترکی دست یابند استفاده می‌شوند. عامل‌ها تلاش می‌کنند با داشتن توابع مشاهده محلی خود که بر اساس اطلاعات آن‌ها از جهان و منابع آن‌ها متفاوت است، بهره مشترک خود را بیشینه کنند. این مدل‌ها در سطح مشاهده عامل‌ها از جهان، میزان تمرکز و نیز ارتباط بین آن‌ها تفاوت دارند [۱۶]. اصلی‌ترین نمونه این مدل‌ها رویکرد DEC-POMDP[1] است که سایر مدل‌ها به نوعی تغییر یافته این رویکرد هستند. مدل DEC-POMDP به دلیل پوشش دادن اکثر پارامترهای محیط‌های فرماندهی و کنترل همچون پشتیبانی از میدان دید محدود، پویایی محیط، عدم قطعیت مشاهده و اقدامات در محیط و نیز طرح‌ریزی غیر متمرکز از اهمیت بالایی برخوردار است. از این رو نسبت به سایر رویکردها بیشتر به آن پرداخته می‌شود. در زیر پس از معرفی رویکرد DEC-POMDP قوی‌ترین توسعه این رویکرد را نیز معرفی می‌شود.

مدل DEC-POMDP: مدل DEC-POMP یک توسعه طبیعی از چارچوب زنجیره تصمیم مارکوف را برای شرایط چند عاملی و مشارکتی پیشنهاد می‌دهد که برای طرح‌ریزی و یادگیری در شرایط عدم قطعیت بسیار مفید است. در این مدل عامل‌ها نمی‌توانند با هم ارتباط داشته باشند. یک فرآیند تصمیم‌گیری مارکوف نیمه رؤیت پذیر غیر متمرکز یک چند تایی <I, S,{Ai}, { Ωi},P, O, R, b0> است که

- I مجموعه متناهی از عامل‌ها با اندیس‌های ۱،... ،n است.
- S یک مجموعه متناهی از حالت‌های سامانه است. یک حالت[2] تمام ویژگی‌های مربوط به سامانه پویا و ویژگی‌های مارکوف را دارا است.
- Ai مجموعه محدودی از اقدامات در دسترس برای عامل i است و A=XiϵI Ai مجموعه اقدامات مشترک است که a=< a1... an > یک اقدام مشترک را نشان می‌دهد.
- Ωi مجموعه متناهی از مشاهدات در دسترس برای عامل i است و Ω =XiϵI Ωi مجموعه مشاهدات است که o=<o1, ... , on > یک مشاهده اشتراکی است.
- P یک جدول احتمال انتقال حالت است. P(s´|s, a) نشان دهنده احتمال اتخاذ اقدام a در جابه‌جایی از حالت s به حالت s´ است.
- O یک جدول احتمالات مشاهدات است. O(o |s´, a) نشان دهنده احتمال دیدن مشاهده اشتراکی o بعد از اقدام مشترک a و رسیدن به حالت s´ است.
- R: S*A →R تابع بهره است. R(s, a) نشان دهنده بهره به‌دست آمده از اتخاذ اقدام مشترک a در حالت s است.
- b0ϵ Δ(S) توزیع حالت گمان اولیه است.

خط مشی محلی برای DEC-POMDP: یک خط مشی محلی برای عامل i، یعنی δi ، نگاشتی از پیشینه‌های محلی مشاهدات ($\overline{O} = O_{i1}, \dots, O_{it}$) به اقدامات Ai است یعنی:

δi= Ωi*→ Ai.

خط مشی مشترک برای DEC-POMDP: یک خط مشی مشترک، δ =< $\delta_1, \dots, \delta_n$ > یک چند تایی از سیاست‌های معین محلی است و هر کدام برای یک عامل است، یعنی، $\delta(h) = < \delta_1(h_1), \dots, \delta_n(h_n) >= \vec{a}$

حل یک DEC-POMDP: منظور از حل، یافتن یک خط مشی مشترک برای بیشینه کردن بهره کلی مورد انتظار است. چنین سیاستی با بیشینه کردن مقدار آن می‌تواند به‌دست آید. یعنی با در نظر گرفتن هر دنباله ممکنی از جوایز و ضرب آن در احتمال آن و سپس میانگین‌گیری آن قابل دست‌یابی است. این مقدار مجموع بهره مورد انتظار آینده است. چون انتقال از یک حالت به حالت دیگر معین نیست، مقدار یک خط مشی مشترک δ برای یک DEC-POMDP با دامنه محدود با حالت اولیه s_0 برابر است با

$$V^{\delta}(s_0) = E\left[\sum_{t=0}^{T-1} R(\vec{a_t}, s_t) | s_0, \delta\right] \qquad (۱)$$

به طور کلی رویکردهای مبتنی بر فرایند تصمیم مارکوف از پیچیدگی محاسباتی بالا رنج می‌برند. این مسئله با افزایش تعداد عامل‌ها و نیز تنوع عامل‌ها شدت بیشتری می‌یابد. مدلی که در ادامه بیان می‌شود قوی‌ترین توسعه مدل DEC-POMDP است.

مدل MAOP-COMM: این الگوریتم یکی از جدیدترین توسعه‌های مدل DEC-POMDP به شمار می‌رود که طرح‌ریزی را

[1] Decentralize Partially Observable Markov Decision Process
[2] State

www.SID.ir

Downloaded from adst.ir at 11:32 +0430 on Saturday August 12th 2017

به صورت برخط انجام می‌دهد. یعنی در هر گام زمانی طرح‌ریزی و اجرا با هم آمیخته می‌شود. به علاوه ارتباطات بین عامل‌ها محدود است یعنی نه ارتباطات آزاد نه بدون ارتباط. رویکرد MAOP-COMM ویژگی‌هایی همچون پویایی، برخط بودن طرح‌ریزی، توانایی مواجه با عدم قطعیت و طرح‌ریزی با وجود محدود بودن ارتباطات را دارا است. در این روش هر عامل به‌صورت مستقل طرح‌ریزی نموده و بر اساس مشاهدات محلی خود انبار باوری می‌سازد و هر گاه مشاهدات جدید وی با انبار باور در تناقض باشد، با عامل همکار خود ارتباط برقرار می‌کند [۳]. انبار باور چنین تعریف می‌شود: یک "انبار باور[1]" در گام زمانی t به‌وسیله چند تایی $< \{H_i^t | i \in I\}, B^t >$ تعریف می‌شود که Hit مجموعه‌ای از پیشینه‌ها برای عامل i است و Bt مجموعه‌ای از حالات باور مشترک است [۳]:

$$Bt = \{(h^t) | h^t \in H^t\}$$

$$H^t = \sum_{i \in I} H_i^t$$

اگر مجموعه‌ای از پیشینه‌های مشترک از گام پیش داشته باشید، با استفاده از قاعده بیز می‌توان مجموعه‌ای از حالت‌های باور مشترک گام جاری ($b^t(s'|h^t)$) را محاسبه کرد [۳]:

$$\forall s' \epsilon S \ , \ b^t(s'|h^t) = \frac{O(o^t|s', a^{t-1}) \sum_{s \epsilon S} P(s'|s, a^{t-1}) b^{t-1}(s|h^{t-1})}{\sum_{s'' \epsilon S} O(o^t|s'', a^{t-1}) \sum_{s \epsilon S} P(s''|s, a^{t-1}) b^{t-1}(s|h^{t-1})} \quad (۲)$$

در بخش سوم بیشتر به این رویکرد پرداخته می‌شود و پس از ذکر معایب آن، با استفاده از دو بهبود پیشنهادی، کارایی الگوریتم آن را افزایش داده می‌شود.

مقایسه روش‌های طرح‌ریزی: مدل‌های نظری تصمیم‌گیری برای طرح‌ریزی چند عاملی را به راه‌های مختلفی می‌توان دسته‌بندی کرد. در جدول (۱) روش‌های طرح‌ریزی یاد شده، از نظر برخی ویژگی‌های محیط‌های فرماندهی و کنترل مقایسه شده است. از میان روش‌های موجود رویکرد MAOP-COMM به دلیل دارا بودن ویژگی‌هایی همچون پویایی، برخط بودن طرح‌ریزی، توانایی مواجه با عدم قطعیت و طرح‌ریزی با وجود محدود بودن ارتباطات برای طرح‌ریزی در محیط فرماندهی و کنترل نسبت به سایر رویکردها مناسب‌تر به نظر می‌رسد. البته این مدل از پیچیدگی زمان و حافظه بالا رنج می‌برد. عامل‌های BDI نیز به دلیل انعطاف بالا، تصمیم‌گیری و استدلال مبتنی بر آگاهی و خودمختاری، برای طرح‌ریزی در محیط‌های مورد نظر مناسب هستند، اما محدودیت‌هایی نیز دارند که به آن‌ها اشاره شده است [۵].

[1] Belief Pool

۲-۳. الگوریتم پیشنهادی برای بهبود کارایی رویکرد MAOP-COM

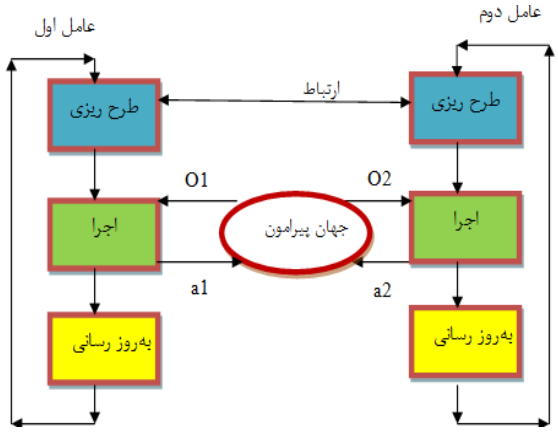
در این بخش ابتدا رویکرد MAOP-COMM را به صورت دقیق‌تر مورد مطالعه قرار داده و نحوه سیاست‌گذاری آن را در طرح‌ریزی بیان می‌شود. سپس نقاط ضعف آن را مورد بررسی قرار داده و در انتها با استفاده از تغییرات پیشنهادی در الگوریتم این مدل، کارایی مدل را افزایش داده می‌شود

نحوه طرح‌ریزی در رویکرد MAOP-COMM: همان‌طور که قبلاً بیان شد، رویکرد MAOP-COMM به دلیل دارا بودن ویژگی‌هایی همچون پویایی، برخط بودن طرح‌ریزی، توانایی مواجه با عدم قطعیت و طرح‌ریزی با وجود محدود بودن ارتباطات، مطابقت زیادی با پارامترهای سناریوهای فرماندهی و کنترل دارد.

الگوریتم MAOP-COMM به صورت موازی به وسیله تمام عامل‌های تیم اجرا می‌شود که در آن طرح‌ریزی و اجرا گنجانده شده است، به طور دقیق‌تر، این الگوریتم بر خط به فاز طرح‌ریزی، فاز اجرا و فاز به‌روزرسانی تقسیم می‌شود که به صورت متوالی در هر گام زمانی اجرا می‌شوند.

در شکل (۲) نمودار مدل MAOP-COMM برای طرح‌ریزی بین دو عامل نشان داده شده است. در الگوریتم MAOP هر عامل به صورت مستقل طرح‌ریزی می‌کند تا اقدام مناسب برای لحظه کنونی را بیابد. عامل اقدام را در فاز اجرا انجام می‌دهد و در نتیجه به حالت جدیدی از سناریو می‌رود. در حالت جدید مشاهده دریافتی از محیط را ذخیره می‌کند تا در طرح‌ریزی گام بعد آن را به‌کار برد. در مرحله بعد باورها به‌روزرسانی می‌شوند و دوباره عامل وارد فاز طرح‌ریزی می‌شود [۱۷].

چنانچه مشاهده دریافتی عامل با باورهایش ناسازگار باشد، عامل اقدام به ارتباط با عامل همکار خود می‌کند. وقتی یک عامل باور دارد که p باید درست باشد و مشاهدات عامل اظهار می‌دارند که ~p، باور آن عامل ناسازگار خوانده می‌شود.



شکل ۲. ساختار طرح‌ریزی بر خط بین دو عامل با ارتباطات محدود (o مشاهده دریافتی از محیط و a اقدام عامل) [۱۷]

Arhive of SID

www.SID.ir

Downloaded from adst.ir at 11:32 +0430 on Saturday August 12th 2017

جدول ۱. دسته‌بندی روش‌ها بر اساس ویژگی‌های مورد نیاز برای محیط‌های فرماندهی و کنترل [۵]

مدل	شناختی	پیچیدگی زمانی	پویایی	عدم قطعیت	عدم تمرکز	برخط	توضیحات
BDI	بله	P-COMPLETE	بله	بله	بله	بله	عدم دارا بودن سازوکار- نیاز به مدل کردن عامل بر طبق سناریو
Game Theory	خیر	P-COMPLETE	خیر	بله	بله	خیر	بر اساس تئوری تصمیم
DCOP	خیر	complex	بله	خیر	بله	خیر	بهینه‌سازی محدودیت
MDP's	خیر	P-COMPLETE	خیر	بله	بله	خیر	نامناسب برای محیط‌های نیمه رؤیت‌پذیر
DEC-POMDP	خیر	NP-COMPLETE	بله	بله	بله	خیر	پیچیدگی زمانی زیاد- برون‌خط
MAOP-COMM	خیر	NP-HARD	بله	بله	بله	بله	رفع معایب DEC-POMDP، افزایش پیچیدگی حافظه و زمان با افزایش تعداد عامل و تنوع اقدامات

در مرحله زمانی t، انبار باور Bt نگه داشته شده، ناسازگار با مشاهده محلی oit گفته می‌شود اگر

$$\max_{\forall b,\forall o^t_{-i}} \{ \sum_{s' \in S} O(O^t|s',\vec{a}) \sum_{s \in S} P(s'|s,\vec{a}) b(s) \} < \varepsilon \qquad (۳)$$

$$b \in B(h^t_i) \quad \overrightarrow{O^t} = o^t_i \cup o^t_{-i}, \qquad o^t_{-i} \in \sum_{k \neq i} \Omega_k$$

که در آن، a یک اقدام مشترک مبتنی بر ot و خط مشی مشترکی است که در گام قبل محاسبه شده است. آستانه ε به وسیله ساختار تابع مشاهده معین می‌شود [۳].

یافتن خط مشی بهینه در MAOP-COMM هدف اصلی هر الگوریتم طرح‌ریزی یافتن خط مشی‌های بهینه‌ای است که اقدامات درست را اتخاذ کنند. هر عامل i یک خط مشی محلی دارد که نگاشتی از پیشینه‌ها به اقدامات است که با δi(hi) نمایش داده می‌شود: δ i(hi) :Hi →Ai i(hi). δi(hi) نشان دهنده اقدامی است که به پیشینه hi منصوب می‌شود. خط مشی‌های همه عامل‌ها با هم تشکیل یک خط مشی مشترک می‌دهد و یک خط مشی مشترک مجموعه‌ای از خط مشی‌های محلی است . . . ,δ=<δ1,δ2 , δn> که هر δi برای یک عامل است و δ(h) نشان دهنده اقدام مشترک متعلق به پیشینه مشترک h است. هر عامل به طور مستقل همان طرح δ(h) را برای تیم محاسبه می‌کند و سپس بر اساس پیشینه محلی خود، بخشی از طرح را که مربوط به او است، اجرا می‌کند. اگر δ(h) خط مشی کل مسیر باشد، q یک خط مشی از زمان جاری به انتهای مسئله است. به منظور یافتن خط مشی qi برای پیشینه hi عامل i، عامل‌ها باید درباره تمام پیشینه‌های ممکن h-i که به وسیله دیگران نگه داشته شده و همچنین تمام خط مشی‌های ممکن مربوط به آن‌ها استدلال کنند. به عبارت دیگر، نیاز است که یک خط مشی مشترک δ بیابید که تابع ارزش زیر را بیشینه کند:

$$V(\delta) = \sum_{h \in H} \sum_{s \in S} P(s|h) V(\delta(h), s) \qquad (۴)$$

یافتن خط مشی‌های غیر متمرکز، شبیه به ساختن درخت خط مشی است که در الگوریتم‌های طرح‌ریزی DEC-POMDP به‌کار می‌روند. همان‌طور که در شکل (۳) نشان داده شده است، پیشینه مشاهدات و اقدامات ($h_1, h_2, \ldots, h_n$)، مسیرهایی از درخت از سمت ریشه به شاخه جاری هستند.

راه مستقیم یافتن بهترین خط مشی مشترک این است که تمام نگاشت‌ها از پیشینه‌ها به زیر خط مشی‌ها به شمار آیند و بهترین آن‌ها انتخاب شود. این مسئله با مسئله تصمیم‌گیری غیر متمرکز که در مرجع [۱۸] آمده است، هم‌ارز است که ثابت شده است که پیچیدگی آن NP-HARD است. از این رو، در الگوریتم ارائه شده، با استفاده از توابع هیوریستیک یک راه‌حل تقریبی یافت شد. این راه‌حل عبارت است از حل مسئله به عنوان یک برنامه خطی. در این راه‌حل، مقدار خط مشی مشترک یعنی π به صورت زیر محاسبه می‌شود [۳]:

$$V(\pi) = \sum_{h \in H} p(h) \sum_{\vec{q}} \prod_{i \in I} \pi_i(q_{i|h_i}) Q(\vec{q}, b(h)) \qquad (۵)$$

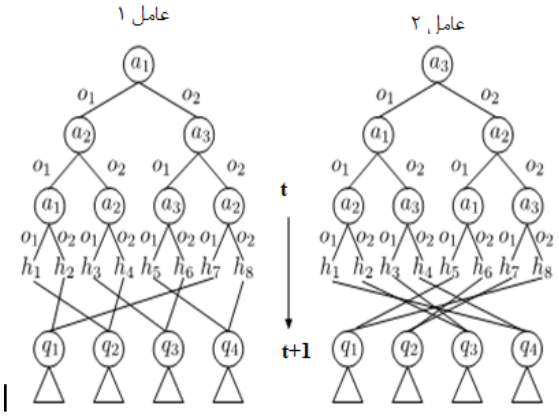
که در آن، p(h) توزیع احتمال پیشینه h است، b(h) حالت باوری است که توسط h استنتاج شده است و $Q(\vec{q}, b(h))$ مقدار خط مشی q در b(h) است. q یک خط مشی از زمان جاری به انتهای مسئله است، بنابراین، $Q(\vec{q}, b(h))$ مقداری است که عامل‌ها در گام‌های آتی هنگامی‌که با یک توزیع حالت b(h) شروع می‌کنند، به آن دست می‌یابند. متأسفانه مقدار بهینه Q به طور برخط در دسترس نیست. اما مقدار بهینه می‌تواند با هیوریستیک معینی تخمین زده شود. معمولاً هیوریستیکی که نزدیک به مقدار بهینه باشد، ممکن است زمان بیشتری برای محاسبه صرف کند اما سبب نتیجه بهتری می‌شود و برعکس. هر روشی که مجموعه‌ای

Arrive of SID

www.SID.ir

Downloaded from adst.ir at 11:32 +0430 on Saturday August 12th 2017

از توابع مقدار V(s) را فراهم کند، می‌تواند به عنوان هیوریستیک به‌کار رود. در حالت ایده‌آل، هیوریستیک نه تنها باید مقدار ناگهانی اقدام مشترک را ارائه دهد، بلکه مقدارهای آتی نیز مورد انتظار هستند.



شکل ۲. یافتن خط مشی‌ها برای طرح‌ریزی بر خط را می‌توان به صورت پیمایش درخت خط مشی نشان داد [۱]

در رویکرد MAOP-COMM پیش‌بینی تک گامی برای محاسبه مقدار گام‌های آینده به‌کار رفته است. بدین معنی که سیاستی با فقط یک گره اقدام در نظر گرفته شده است. برای پیش‌بینی تک گام، q_i و q می‌توانند با a_i و a جایگزین شوند. برای تخمین مقدار بهینه Q، حل MDP اساسی به عنوان هیوریستیک به‌کار رفته است. هیوریستیک Q_{MDP} می‌تواند به صورت زیر نوشته شود:

$$Q(\vec{a},b)=\sum_{s\in S}b(s)\left[R(s,\vec{a})+\sum_{s'\in S}P(s'|s,\vec{a})V_{MDP}(s')\right] \qquad (۶)$$

که در آن، V_{MDP} تابع مقدار MDP اصلی است. هیوریستیک Q_{MDP} حد بالای مقدار بهینه است. چون بر اساس این فرضیه است که عامل‌ها حالت سامانه مورد نظر را در هر گام به طور کامل می‌بینند. درنهایت به دنبال خط مشی π هستیم که مقدار $V(\pi)$ را بیشینه کرده و نامساوی زیر برقرار باشد:

$$V(\pi)\le\sum_{h\in H}P(h)\left[\sum_{\vec{q}}\pi_i(q_i|h_i)\pi_{-i}(q_{-i}|h_{-i})Q(\vec{q},b(h))\right] \qquad (۷)$$

خط مشی مشترکی که V(π) را بیشینه کند، خط مشی منتخب خواهد بود و اقدام پیشنهادی آن توسط عامل‌ها اجرا می‌شود [۳].

بهبود رویکرد MAOP-COMM با الگوریتم پیشنهادی: الگوریتم MAOP-COMM اگر چه برای محیط‌های پویا با میدان دید محدود و نیازمند طرح‌ریزی برخط طراحی شده است، اما در شرایط بحرانی محیط‌های فرماندهی و کنترل که نیاز به تصمیم‌گیری دقیق، ولو با صرف زمان بیشتر ارجحیت دارد، کارایی مطلوبی ندارد. یکی از علل آن این است که این الگوریتم برای یافتن خط مشی بهینه در درخت خط مشی تشکیل شده از پیشینه‌ها، از پیش‌بینی تک گام استفاده می‌کند. بدین معنی که درخت سیاستی با فقط یک گره اقدام در نظر گرفته می‌شود. این امر منجر به کاهش دست‌یابی به سیاست‌های دقیق و اتخاذ تصمیمات نادرست می‌شود. علت دیگر آن این است که این الگوریتم برای تخمین گام‌های آتی اقدامات عامل‌ها (پیش‌بینی گره‌های بعد از پیش‌بینی تک گام) از هیوریستیک MDP استفاده کرده است، در حالی که دامنه، یک مسئله DEC-POMDP است. یعنی مسئله تصمیم‌گیری با اطلاعات نیمه رؤیت پذیر بین چند عامل که اقدامات آن‌ها بر یکدیگر اثرگذار است را به یک مسئله تصمیم‌گیری تک عاملی با رؤیت پذیری کامل تقریب می‌زند و این امر منجر به تخمین غیر دقیق از پاسخ مسئله می‌شود [۱۹].

به‌کارگیری هیوریستیک پیشنهادی برای یافتن سیاست‌های تقریباً بهینه: در طرح‌ریزی برخط یافتن مقدار بهینه سیاست‌ها دشوار است. چون برای حل آن باید به اندازه حل تمام DEC-POMDP کار کرد. مقدار بهینه می‌تواند با هیوریستیک معینی تخمین زده شود. در حالت ایده‌آل، هیوریستیک نه تنها باید مقدار ناگهانی اقدام مشترک را ارائه دهد، بلکه مقدارهای آتی نیز مورد انتظار هستند. رویکردی که در روش‌های موجود استفاده شده است، به‌کارگیری هیوریستیک MDP است. معمولاً هیوریستیکی که نزدیک به مقدار بهینه باشد، ممکن است زمان بیشتری برای محاسبه صرف کند اما سبب نتیجه بهتری می‌شود و برعکس. بنابراین هیورستیک POMDP را برای تخمین گام‌های آتی به‌کار می‌رود که این هیوریستیک به صورت زیر است:

$$Q(\vec{a},b)=\sum_{s\in S}b(s)\left[R(s,\vec{a})+\sum_{s'\in S}P(s'|s,\vec{a})\sum_{\vec{o}\in\vec{\Omega}}O(\vec{o}|s',\vec{a})\,V_{POMDP}(b_{\vec{a}}^{\vec{o}})\right] \qquad (۸)$$

که در آن، $b_{\vec{a}}^{\vec{o}}$ حالت باور جانشین b با $\vec{a}$ و $\vec{o}$ است و V_{POMDP} تابع مقدار POMDP مورد نظر است [۲۰]. به طور شهودی، به‌کارگیری این هیوریستیک به آن معنا است که عامل‌ها مشاهدات خود را در هر گام بعدی با هم به اشتراک می‌گذارند.

$$V_{POMDP}=R(s)+\gamma\max(a')\sum p(s'|s,a)V(s') \qquad (۹)$$

رابطه فوق برای طرح‌ریزی یک عامل در محیط غیر قطعی و نیمه رؤیت پذیر به‌کار می‌رود [۲۰]. ما از این رابطه برای محاسبه هیوریستیک POMDP استفاده می‌شود.

روش پیشنهادی برای یافتن دقیق‌تر سیاست‌ها: اتخاذ پیش‌بینی چند گام در درخت خط مشی برای پیش‌بینی تک گام $V(s')$ را برابر R(s′) (بهره دریافتی از محیط در حالت s') قرار

Arhive of SID

www.SID.ir

Downloaded from adst.ir at 11:32 +0430 on Saturday August 12th 2017

داده می‌شود و برای پیش‌بینی تا چند گام فراتر و یافتن دقیق‌تر سیاست، $V(s')$ را توسعه داده می‌شود. کارهای انجام شده پیش‌بینی تک گامی را برای یافتن اقدام بهینه به‌کار برده‌اند. بدین معنی که درخت سیاستی با فقط یک گره اقدام در نظر می‌گرفتند. این امر منجر به کاهش دست‌یابی به سیاست‌های دقیق و اتخاذ تصمیمات نادرست می‌شد. در اینجا پیش‌بینی دو گام را به‌کار می‌رود تا خط مشی دقیق‌تری در طرح‌ریزی اتخاذ شود. بدین معنی که درخت خط مشی نشان داده شده در شکل (۳) را تا عمق دوم توسعه داده می‌شود و از رابطه (۱۰)، برای توسعه تا عمق دوم استفاده می‌شود.

$$V(s') = R(s') + \gamma \max a' \sum_{s'} P(s''|s',a')V(s'') \qquad (۱۰)$$

برای شروع فرآیند جستجو، هر خط مشی محلی π_i با انتخاب یک اقدام تصادفی (a_i) با توزیع یکنواخت مقداردهی می‌شود تا معین شود. سپس با استفاده از برنامه‌نویسی خطی [1] بهترین خط مشی ممکن را که بهترین اقدامات را به‌دست دهد، یافت می‌شود. به این صورت که هر عامل به صورت دوره‌ای انتخاب می‌شود و در حالی که خط مشی دیگر عامل‌ها ثابت در نظر گرفته می‌شود، خط مشی آن عامل بهبود داده می‌شود. برای تحقق این امر برای عامل iام بهترین اقدام a_i^* (و خط مشی مربوط به آن q_i) را در مجموعه h_iها یافت می‌شود که $V(\pi)$ در فرمول (۷) بیشینه شود [۱۹]. در الگوریتم (۱)، رویه اتخاذ خط مشی و اقدام بهینه با استفاده از هیوریستیک POMDP و پیش‌بینی دو مرحله‌ای نشان داده شده است.

الگوریتم ۱. الگوریتم پیشنهادی اتخاذ خط مشی بهینه

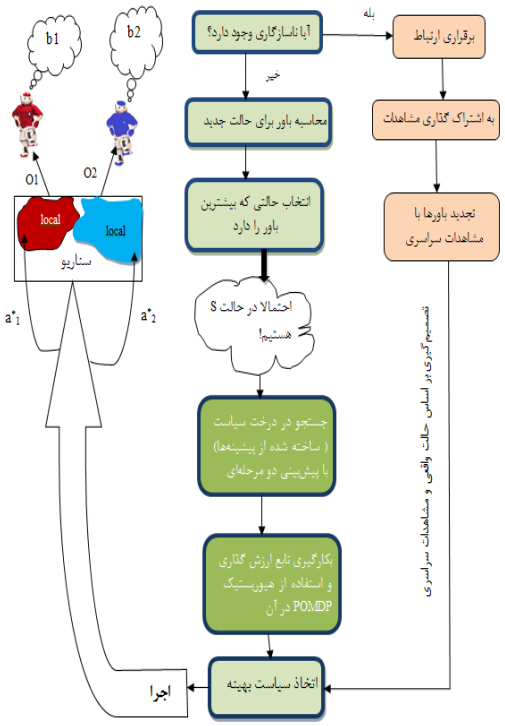
Input:
current state
O_i^t :local Observation
h_i^t : local history
b^t : Beleif of current state

Output:
a^*: Actions for current state

1. for all s′ = nextState (s,$\vec{a}$) // going to state s' by doing action $\vec{a}$
2. for all s'' = nextState (s′, $\vec{a'}$)
3. $Q(\vec{a'}, b^t) = \sum_{s \in S} b\,[R(s', \vec{a'}) + \sum_{o \in O} P(o|s'', \vec{a'})V_{pomdp}$
4. a*= argmax $_{\vec{a}}$ {[$R(s,\vec{a})$P(s′| s,$\vec{a}$) + R(s′,a′)P(s''|s′,a′)]b^t + $Q(\vec{a'}, b^t)$}
5. return a^*

عامل‌ها پس از دریافت مشاهدات محلی خود از محیط (در صورتی که مشاهده دریافتی با باور فعلی در تناقض نباشد)، بر اساس باوری که از حالت کنونی دارند، باوری از حالت جدید می‌سازند. در حالت جدید با ایجاد خط مشی بهینه، بهترین اقدام را برای خود و همکاران محاسبه می‌کنند و سپس سهم خود یعنی اقدام محلی خود را از آن خط مشی اجرا می‌کنند. نمودار (شکل (۴))، نحوه عملکرد رویکرد طرح‌ریزی مشارکتی چند عاملی مبتنی بر مدل MAOP-COMM با بهبودهای پیشنهادی را به صورت مشروح نشان می‌دهد.

پیچیدگی الگوریتم حاصل: الگوریتم‌های DECPOMDP دارای پیچیدگی زمانی NP-HARD هستند. رویکرد MAOP-COMM نیز که زیرمجموعه رویکردهای DEC-POMDP است، همین پیچیدگی زمانی را دارد. بهبودهای پیشنهادی علی‌رغم بهبود کارایی طرح‌ریزی به دلیل بسط درخت تصمیم تا دو مرحله برای هر یک از اقدامات عامل‌ها این پیچیدگی را افزایش می‌دهند. بنابراین با افزایش تعداد عامل‌ها این پیچیدگی به طور نمایی مضاعف می‌شود. این مسئله را در بخش نتایج آزمایش‌ها برای سناریو Grid Soccer خواهید دید.



شکل ۳. نمودار رویکرد طرح‌ریزی مبتنی بر MAOP-COMM با الگوریتم و بهبودهای پیشنهادی.

۲-۴. ارائه رویکرد طرح‌ریزی در قالب وب سرویس

معماری سرویس‌گرا: معماری سرویس گرا یک چارچوب راهبردی از فناوری است که به تمام سامانه‌های داخل و خارج سامانه اجازه ارائه یا دریافت سرویس‌های خوش‌تعریف می‌دهد. این معماری روشی برای طراحی و پیاده‌سازی نرم‌افزارهای گسترده سازمانی به وسیله ارتباط بین سرویس‌هایی است که

[1] Linear Programming

www.SID.ir

Downloaded from adst.ir at 11:32 +0430 on Saturday August 12th 2017

دارای خواص اتصال سست و استفاده مجدد هستند [۲۱]. منظور از سرویس، کاری است که به وسیله یک سرویس دهنده انجام می‌شود. برای پیاده‌سازی ساده یک سرویس می‌توان یک وب سرویس را به‌کار برد که در واقع عینیت یافته سرویس است.

وب سرویس: وب سرویس یک نظام نرم‌افزاری است که برای پشتیبانی قابلیت تعامل دو سامانه مستقل بر پایه شبکه طراحی شده است. وب سرویس‌ها می‌توانند بر روی شبکه و از راه دور قابل دسترسی باشند. سرویس وب نوعی سامانه نرم‌افزاری محسوب می‌شود که برای تعامل ماشین با ماشین در سطح شبکه طراحی شده و دارای یک تعریف قابل پردازش توسط ماشین با نام WSDL است. دیگر سامانه‌ها بر طبق این تعریف از قبل مهیا شده، با سرویس دهنده تعامل خواهند داشت. پیام‌ها توسط پروتکل SOAP (ترکیب HTTP و XML) یا سایر پروتکل‌های مربوطه منتقل می‌شوند.

با به‌کارگیری توانایی سامانه‌های چند عاملی و قابلیت‌های قوی سرویس‌های وب در تعاملات، رویکرد طرح‌ریزی ارتقاء یافته در بخش قبل را با استفاده از فناوری JADE در قالب یک وب سرویس بیان می‌شود. استانداردهایی نظیر FIPA جهت پیاده‌سازی سامانه‌های چند عاملی ارائه شده است و این طور به نظر می‌رسد که چارچوب JADE توانسته است کامل‌ترین چارچوب جهت پیاده‌سازی استاندارد FIPA باشد [۲۲].

برای طراحی وب سرویس بسته به نوع معماری مورد نظرمان برای طرح‌ریزی عامل‌ها از پارادایم AOP[1] یا از استاندارد CXF[2] همراه با به‌کارگیری موتور AXIS[3] برای پروتکل ارتباطی SOAP استفاده می‌شود. اگر عامل را درون بستر JADE در معماری سرویس طرح‌ریزی تعریف شود، به گونه‌ای که سرویس خود، عامل‌های نرم‌افزاری را تعریف کرده و سناریو دریافتی از سرویس گیرنده را برای آن‌ها شبیه‌سازی کند از پارادایم AOP استفاده می‌شود و چنانچه بخواهید عامل‌ها خارج از بستر سرویس چند عاملی باشند به گونه‌ای که بتوان یک کاربر انسانی را جایگزین عامل‌های نرم‌افزاری کرد، می‌توان از استاندارد CXF یا موتور AXIS بهره برد. البته چارچوب JADE قابلیت پیاده‌سازی هر دو معماری را دارد [۲۲].

سرویس طرح‌ریزی طراحی شده می‌تواند برای هر سناریوی دلخواه که مورد نظر کاربر آن باشد طرح‌ریزی نموده و به صورت برخط دنباله اقداماتی را که هر یک از عامل‌ها در هر گام زمانی باید انجام دهند به کاربر نشان دهد. این سرویس در محیط‌های پویا، دارای عدم قطعیت، با میدان دید محدود، دارای عامل‌های توزیع شده و متنوع، به صورت غیر متمرکز قابلیت طرح‌ریزی کارا دارد. به عبارت دیگر پارامترهای دفاع آرایه‌ای (توزیع‌شدگی[4]، عدم تمرکز[5]، پویایی[6]، تنوع[7]) در این سرویس گنجانده شده است. منظور از پشتیبانی از تنوع عامل‌ها آن است که عامل‌ها می‌توانند وظایف متفاوت از حیث نوع و تعداد داشته باشند. این سرویس طرح‌ریزی چون به هر یک از عامل‌ها اجازه طرح‌ریزی می‌دهد، به آن‌ها این امکان را می‌دهد تا در طرح‌ریزی برای خود یا در طرح‌ریزی فرضی برای همکار خود، برای هر عامل اقدامات خاص او را لحاظ کنند. کاربر خارجی جهت استفاده از سرویس طرح‌ریزی برای هر سناریوی مورد نظر، باید آن سناریو را طراحی نموده، و به عنوان ورودی، آن را در اختیار سرویس طرح‌ریزی قرار دهد.

۳. نتایج و بحث

برای اثبات تمامیت سرویس ارائه شده، سرویس ارائه شده با الگوریتم پیشنهادی را روی مسئله مهم Grid Soccer که از جمله جدیدترین سناریوهای ارائه شده برای آزمایش الگوریتم‌های DEC-POMDP است [۳]، آزمایش شد. البته مسائل آزمایشی ساده دیگری نیز برای طرح‌ریزی مشارکتی وجود دارد که اکثراً برای محیط‌های ایستا طراحی شده‌اند. از جمله این مسائل Meeting in a Grid [۲۳] و Cooperative Box-Pushing [۲۴] را نام برد. ساختار این مسائل آزمایشی برای ابزار MADP طراحی شده است. ابزار MADP برای طرح‌ریزی مشارکتی مبتنی بر مارکوف در سامانه عامل لینوکس نوشته شده است [۲۵]. سناریو Grid Soccer نسبت به سناریوهای مذکور، از تمامیت و سطح پیچیدگی بیشتری برخوردار است به طوری که تعداد حالات سناریو آن بیش از ۲۰ برابر سایر سناریوها است. از این رو این سناریو را محور ارزیابی رویکرد ارائه شده قرار داده شد. نتایج شهودی این اجرا بر سناریو Grid Soccer (فوتبال با زمین بازی ۳×۳) برای گام اول و آخر این سناریو طی ۸ گام زمانی در بازی در بخش پیوست، شکل (۵- الف و ب) نشان داده شده است. روش ارائه شده با استفاده از زبان جاوا با پردازنده ۲/۷ GHz گیگاهرتز و حافظه RAM به میزان 4 GB پیاده‌سازی شده است.

تحلیل نتایج الگوریتم طرح‌ریزی با بهبودهای پیشنهادی: الگوریتم رویکرد MAOP-COMM با ارتقاء هیوریستیک تابع ارزش و به‌کارگیری پیش‌بینی دو مرحله‌ای در استراتژی یافتن خط مشی بهینه و اخذ تصمیم صحیح، بهبود یافت و از آن به عنوان هسته مرکزی سرویس مورد نظر استفاده شد. نتایجی که

[1] Aspect Oriented Programming

[2] چارچوبی متن باز است که با استفاده از واسط‌های برنامه نویسی به ساخت و توسعه سرویس‌ها کمک می‌کند.

[3] هسته مرکزی یک سرویس وب

[4] Distributed

[5] Decentralize

[6] Dynamic

[7] Diversity

Arhive of SID

www.SID.ir

Downloaded from adst.ir at 11:32 +0430 on Saturday August 12th 2017

در ادامه آورده می‌شوند نشان می‌دهد الگوریتم حاصل علاوه بر اینکه پارامترهای محیط‌های فرماندهی و کنترل را در خود لحاظ کرده، از کارایی مضاعفی نیز برخوردار بوده و طرح‌ریزی بسیار دقیق‌تری نسبت به الگوریتم‌های اخیر انجام می‌دهد.

در جدول (۲)، کارایی دو الگوریتم با ارتباطات محدود بر روی دو دامنه (فوتبال ۲×۳ و ۳×۳) مورد بررسی قرار گرفته است و بهره تجمعی میانگین (بهره)، میانگین زمان اجرای بر خط در هر مرحله (زمان (s)) و میانگین درصد ارتباط بین عامل‌ها در طول مدت طرح‌ریزی (ارتباط (%)) با هر دو دامنه با روش ارائه شده در مقایسه با بهترین روش پیشین یعنی MAOP-COMM ارائه شده است.

همان‌طور که در جدول‌ها قابل مشاهده است، مدت زمان پاسخگویی روش ارائه شده از روش پیشین بیشتر است، اما میزان بهره دریافتی آن (رابطه (۷))، که نشانگر کارایی طرح‌ریزی انجام شده است به نحو چشمگیری بیشتر است. اگر چه هر دو روش موجود در جدول دارای پیچیدگی زمانی NP-HARD هستند، علت افزایش زمان پاسخگویی در الگوریتم ما، استفاده از پیش‌بینی دو مرحله‌ای به جای پیش‌بینی تک گام است. یعنی برای هر حالت ممکن از فضای مسئله درخت خط مشی را تا عمق دوم توسعه می‌دهد. اما روش MAOP پیشین پیش‌بینی تک گام را برای یافتن خط مشی اتخاذ کرده است. به علاوه هیوریستیک POMDP نیز به دلیل به‌کارگیری توزیع احتمال حالات ممکن بعدی، از سرعت اجرا می‌کاهد.

تأثیر ارتباط بر میزان کارایی هر دو الگوریتم طرح‌ریزی در جدول (۳) برای سناریو فوتبال ۲×۳ و در جدول (۴) برای سناریو ۳×۳ نشان داده شده است. نتایج به‌دست آمده نشان می‌دهد، وقتی عامل‌ها هیچ ارتباطی با هم ندارند باز هم بهره دریافتی از محیط مطلوب است و این امر نشان می‌دهد که طرح‌ریزی در شرایطی که کانال ارتباطی قطع باشد نیز به خوبی طرح‌ریزی می‌کند.

البته همین الگوریتم در حالتی که ارتباطات آزاد و دائمی برقرار است، کارایی بیشتری از حالت بدون ارتباط دارد. چرا که عامل‌ها مشاهدات خود را دائماً با هم به اشتراک می‌گذارند و میدان دید محدود نیست. اختلاف بهره تجمعی در رویکرد با ارتباط و بدون ارتباط در سناریو ۳×۳ بیشتر است. به نظر می‌رسد در سناریوهای بزرگ‌تر تأثیر ارتباطات آزاد بیشتر باشد، به بیان دیگر، هر چه سناریو بزرگ‌تر باشد، ارتباطات بیشتر، موجب می‌شود که عامل‌ها مشاهدات کامل‌تری داشته باشند و در نتیجه بازده بیشتری در طرح‌ریزی به‌دست آورند.

جدول ۲. مقایسه کارایی روش ارائه شده و MAOP-COMM

الگوریتم	بهره	مدت زمان (s)	ارتباط (٪)	تعداد دفعات
MAOP-COMM ۳×۳	۲۹۶	۲/۳	۲۷	۲۰
	۱۶۷۹/۵	۲/۵	۲۶/۹	۱۰۰
روش پیشنهادی ۳×۳	۷۱۰/۸	۱۲	۲۳	۲۰
	۳۵۵۸/۲	۴۱	۲۴	۱۰۰
MAOP-COMM ۲×۳	۲۹۰/۶	۰/۲۸	۱۴/۸	۲۰
	۱۹۳۳/۹	۰/۱۶	۱۵/۴	۱۰۰
روش پیشنهادی ۲×۳	۱۴۲۴/۳	۱۰	۱۰/۴	۲۰
	۳۶۶۸/۴	۳۹	۱۲	۱۰۰

جدول ۳. مقایسه تأثیر ارتباط در روش ارائه شده و MAOP (برای فوتبال ۲×۳ تعداد حالات ۳۸۴۳، تعداد مشاهدات ۱۱، تعداد اقدامات ۶)

الگوریتم	بهره	مدت زمان (s)	ارتباط (٪)	تعداد دفعات
MAOP	۱۸۰/۵	۰/۲۵	۰	۲۰
	۱۱۵۷/۸	۰/۱۴	۰	۱۰۰
روش پیشنهادی- بدون ارتباط	۶۷۰/۲	۱۳	۰	۲۰
	۳۱۶۱/۲	۳۸	۰	۱۰۰
MAOP FULL-COMM	۳۷۳/۹	۰،۰۱	۱۰۰	۲۰
	۱۹۳۳/۶	۰،۰۱	۱۰۰	۱۰۰
روش پیشنهادی- ارتباطات آزاد	۷۰۹/۲	۱۱	۱۰۰	۲۰
	۳۵۵۹/۶	۳۸	۱۰۰	۱۰۰

Arhive of SID

www.SID.ir

Downloaded from adst.ir at 11:32 +0430 on Saturday August 12th 2017

جدول ۴. مقایسه تأثیر ارتباط در روش پیشنهادی و MAOP (برای فوتبال ۳×۳ تعداد حالات ۱۶۱۳۱، تعداد مشاهدات ۱۱، تعداد اقدامات۶)

الگوریتم	بهره	مدت زمان (s)	ارتباط (٪)	تعداد دفعات
MAOP	۱۹۰٫۷	۱/۹	۰	۲۰
	۸۰۳/۶	۱/۹۶	۰	۱۰۰
روش پیشنهادی- بدون ارتباط	۶۷۰	۱۴	۰	۲۰
	۲۹۵۳/۳	۴۲	۰	۱۰۰
MAOP FULL-COMM	۳۵۶	۰/۰۱	۱۰۰	۲۰
	۱۸۰۸/۲۰	۰/۰۱	۱۰۰	۱۰۰
روش پیشنهادی- ارتباطات آزاد	۷۱۹/۲	۱۲	۱۰۰	۲۰
	۳۵۴۹/۲	۴۲	۱۰۰	۱۰۰

نحوه به‌کارگیری سرویس طرح‌ریزی ارائه‌شده: سرویس طرح‌ریزی طراحی شده می‌تواند برای هر سناریوی دلخواه که مورد نظر کاربر آن باشد طرح‌ریزی نموده و به صورت برخط دنباله اقداماتی را که هر یک از عامل‌ها در هر گام زمانی باید انجام دهند، به کاربر نشان می‌دهد به شرط آنکه کاربر سناریوی خود را به گونه‌ای شبیه‌سازی نماید که ورودی‌های مورد نیاز سرویس طرح‌ریزی را برای آن فراهم نماید. این ورودی‌ها عبارت‌اند از: تعداد عامل‌های سناریوی مورد کاربرد، تعداد اقدامات و مشاهدات پیرامونی ممکن، احتمال هر اقدام و مشاهده، تعداد گام زمانی دلخواه، حالات اولیه، خاص و هدف سناریو، فضای حالت (کاربر باید برای سناریوی خود تابعی را کدنویسی کند که همه حالات ممکن (مثلاً در قالب یک Map یا جدول هش[1]) را تولید نماید)، تابع احتمال قطعیت یک مشاهده، تابع احتمال انتقال از یک حالت به حالت دیگر و غیره. بدین ترتیب سرویس طرح‌ریزی ارائه شده را می‌توان در بسیاری از دامنه‌های حوزه فرماندهی مانند لجستیک، توپخانه و غیره به‌کار برد.

محدودیت‌های سرویس طرح‌ریزی ارائه‌شده: سرویس طرح‌ریزی ارائه شده علی‌رغم قابلیت بالا در انجام طرح‌ریزی با دقت زیاد، برای هر نوع سناریو گسترده، قابل پیاده‌سازی نیست. برای کار با این سرویس باید به محدودیت‌های آن توجه نموده و پارامترهای سناریو مورد آزمون را منطبق با توانایی‌های آن اتخاذ نمود:

۱. در مدل به‌کار رفته، فضای حالت، تابعی نمایی از تعداد عامل‌ها و فضای مسئله است و تعداد حالات متجاوز از چندین هزار خواهد بود. بنابراین هر چه فضای مسئله بزرگ‌تر شود، سرویس طرح‌ریزی نیازمند حافظه بیشتر بوده و زمان بیشتری را نیز صرف می‌کند. البته در کارهای آتی می‌توان با به‌کارگیری روش‌هایی برای هرس درخت خط مشی و جلوگیری از پردازش‌های تکراری مصرف زمان و حافظه را کاهش داد.

۲. هر چه تعداد عامل‌ها افزایش می‌یابد، تعداد مشاهدات محلی مستقل از هم بیشتر شده، احتمال ناسازگاری باورها بیشتر، تصمیم‌گیری دشوارتر و در نتیجه نیاز به ارتباط بیشتر می‌شود. چنانچه کانال ارتباط توانایی پشتیبانی هم‌زمان از همه عامل‌ها را نداشته باشد، بازده کاهش می‌یابد.

۴. نتیجه‌گیری

در این مقاله ابتدا ویژگی‌های بارز محیط‌های فرماندهی و کنترل (که برای طرح‌ریزی در این محیط‌ها باید مدنظر قرار گیرند)، بیان شدند. در ادامه رویکردهای طرح‌ریزی موجود از قبیل DCOP، عامل‌های BDI، تئوری بازی‌ها و مدل‌های مبتنی بر فرایند مارکوف از حیث پارامترهای مختلف محیط‌های فرماندهی و کنترل مورد بررسی قرار گرفتند. از بین رویکردهای معرفی شده جدیدترین توسعه برخط رویکردهای مبتنی بر فرایند تصمیم مارکوف نیمه رؤیت پذیر، یعنی رویکرد MAOP-COMM به دلیل پشتیبانی از پارامترهای محیط‌های فرماندهی و کنترل مورد توجه بیشتری قرار گرفت. برای افزایش کارایی رویکرد ذکر شده، با استفاده از پیش‌بینی دو گام به جای پیش‌بینی تک گام و هیوریستیک POMDP به جای MDP الگوریتم مدل را بهبود داده شد. سپس رویکرد بهبود یافته را در قالب یک وب سرویس پیاده‌سازی شد تا برای هر سناریوی دلخواه با شرایط محدود شده، سناریو شبیه‌سازی شده را در قالب ورودی، از معمار سناریو دریافت کند و با فرض پویا بودن محیط، طرح‌ریزی لازم را به طور برخط انجام داده و دنباله اقدامات عامل‌ها را در هر لحظه ایجاد نماید. نتایج آزمایش‌ها نشان می‌دهد اگر چه مدت زمان طرح‌ریزی الگوریتم جدید نسبت به الگوریتم‌های قبلی به دلیل به‌کارگیری پیش‌بینی دو مرحله‌ای و هیوریستیک قوی‌تر، بیشتر است، میانگین بهره تجمعی حاصل از طرح‌ریزی بسیار بهتر از الگوریتم‌های پیشین است. سرویس طرح‌ریزی طراحی شده می‌تواند برای هر سناریوی دلخواه که مورد نظر کاربر آن باشد طرح‌ریزی نماید، به شرط آنکه کاربر سناریوی خود را به گونه‌ای شبیه‌سازی و توصیف نماید که ورودی‌های مورد نیاز سرویس طرح‌ریزی را برای آن فراهم نماید.

[1] Hash Table

Arhive of SID

www.SID.ir

Downloaded from adst.ir at 11:32 +0430 on Saturday August 12th 2017

۵. مراجع

[1] Russell, S. J.; Norvig, P. "Artificial Intelligence a Modern Approach"; Third Ed., Prentice Hall, 2004.

[2] Yan, Q. "Approach to Multi Agent Planning"; Department of Computer Science, Arizona State University Reports, 2004.

[3] Wu, F.; Zilberstein, S.; Chen, X. "Online Planning for Multi-Agent Systems with Bounded Communication"; Artificial Intelligence, 2011, 175, 487–511.

[4] Easley, D.; Kleinberg, J. "Networks, Crowds, and Markets, Reasoning about Highly Connected World"; Harvard University Press, Cambridge, 2010.

[5] Sadati, S. ; Naghian Fesharaki, M.; Hosseini, S. M. "Survey of Planning Parameters for Command and Control Environments in Existing Planning Models"; 7th Conf. on Command and Control (C4I), Iran, 2013 (In Persian).

[6] Myerson, R. B. "Game Theory: Analysis of Conflict"; Harvard University Press, London, 1991.

[7] Liang, X.; Peng, Y.; Shen, G. Q. "A Game Theory Based Analysis of Decision Making for Green Retrofit under Different Occupancy Types"; Elsevier, 2016, 137, 1300–1312.

[8] Bratman, M. E. "Planning in BDI Agents: A Survey of the Integration of Planning Algorithms and Agent Reasoning"; The Knowledge Engineering Review, 2013, 30, 1–44.

[9] Georgeff, M..; Pell, B.; Pollack, M.; Wooldridge, M. "The Belief-Desire-Intention Model of Agency"; Intelligent Agents V: Agents Theories, Architectures, and Languages 1999, 1555, 1–10.

[10] Hernández, G.; Fallah-Seghrouchni, A. A. E.; Soldano, H. "Learning in BDI Multi-agent Systems"; Proc. of Computational Logic in Multi-Agent Systems, 2004.

[11] Phung, T.; Winikoff, M.; Padgham, L. "Learning within the BDI Framework: An Empirical Analysis"; Knowledge-Based Intelligent Information and Engineering Systems, 2005, 3683, 282-288.

[12] Pokahr, A.; Braubach, L.; Lamersdorf, W. "Jadex: A BDI Reasoning Engine"; MuliAgent Programming, 2008, 15, 149–174.

[13] Sebastian, S.; Silva, L.; Padgham, L. "Hierarchical Planning in BDI Agent Programming Languages: A Formal Approach"; Proc. of the Fifth Int. Joint Conf. on Autonomous Agents and Multi Agent Syst. 2005.

[14] Faltings, B. "Distributed Constraint Programming"; Handbook of Constraint Programming, Elsevier, 2006, 699-729.

[15] Kruk, S. "Planning with Multiple Agents"; Seminar on Autonomous Learning System, 2013.

[16] Seuken, S.; Zilberstein, S. "Formal Models and Algorithms for Decentralized Decision Making Under Uncertainty"; J. of Autonomous Agents and Multi-Agent Syst. 2008, 17, 190–250.

[17] Sadati, S. "Design a Collaborative Online Planning Service based on Markov Process in Command and Control Domain"; M.Sc. Thesis, Malek Ashtar University, Tehran, 2014 (In Persian).

[18] Amato, C.; Zilberstein, S. "Achieving Goals in Decentralized POMDPs"; Proc. of the 8th Int. Joint Conf. on Autonomous Agents and Multi-Agent Syst. 2009, 593–600.

[19] Sadati, S.; NaghianFesharaki, M.; Momeni Azandaryani, M. H. "Online Collaborative Planning in Command and Control Domain"; 7th Conf. on Command and Control (C4I), Iran, 2013 (In Persian).

[20] Weerdt, M.; Mors, A. "Multi-Agent Planning: An Introduction to Planning and Coordination"; Dept. of Software Technology Reports, Delft University of Technology, 2005.

[21] Shams, F.; Mahjourian, A. "Introducing Principles and Methods of Service-Oriented Enterprise Architecture"; Beheshti University Press, Tehran, 2010 (In Persian).

[22] Shooshtarian M.J.; Mohsenzadeh, M. "Provide an Architecture for Service Evaluation and Selection based Multi Agent Systems"; 1th National Conf. on Software Eng. Tehran, 2011, 381-388 (In Persian).

[23] Amato, C.; Bernstein, D.; Zilberstein, S. "Optimizing Fixed-Size Stochastic Controllers for POMDPs and Decentralized POMDPs"; Autonomous Agents and Multi-Agent Syst. 2010, 21, 293–320.

[24] Seuken, S.; Zilberstein, S. "Improved Memory-Bounded Dynamic Programming for Decentralized POMDPs"; Proc. of the 23th Conf. on Uncertainty in Artificial Intelligence, 2012.

[25] Oliehoek, F.; Spaan, M. "MADP Toolbox 0.2"; Technical Report, Informatics Institute, Amsterdam University, 2009.

Arhive of SID

www.SID.ir