دانشگاه صنعتی مالک اشتر
انجمن سیستم های فازی ایران
Intelligent Systems
Scientific Society of Iran
شماره: .....
تاریخ: .....
پیوست: .....
دومین کنگره مشترک سیستم های فازی و هوشمند ایران
۷-۹ آبان ماه ۱۳۸۷
2nd Joint Congress on Fuzzy and Intelligent Systems
28-30 October 2008

# مجرد سازی دانش و انتقال آن با استفاده از یادگیری تقویتی

نرجس زارع[۱] zare.narjes@gmail.com
مجید نیلی احمدآبادی[۲] mnili@ut.ac.ir
احمدرضا ولی[۱] ar.vali@gmail.com
مریم سادات میریان[۲] mmirian@ut.ac.ir

۱- دانشگاه صنعتی مالک اشتر، مجتمع برق و الکترونیک ۲- دانشگاه تهران، دانشکده برق و کامپیوتر، آزمایشگاه رباتیک

**چکیده – هنوز فاصله قابل توجهی بین قابلیت یادگیری در سیستم‌های هوشمند و سیستم‌های بیولوژیکی دیده می‌شود. یکی از دلایل این مساله استفاده نکردن از دانش بدست آمده در طول یادگیری در یک سیستم هوشمند می‌باشد. به منظور نشان دادن کار آمدی این دیدگاه در این مقاله روشی برای انتقال یادگیری ارائه شده است. در این روش دانش یاد گرفته شده در کار مرجع بصورت مفاهیم مجرد در آمده و به کار هدف منتقل می‌شود. عامل با استفاده از این دانش منتقل شده می‌تواند به یادگیری سریعتری در کار هدف دست یابد. نتایج شبیه‌سازی نشان می‌دهند که روش ارائه شده باعث افزایش پاداش بدست آمده درطول یادگیری، بخصوص اوایل یادگیری و در نتیجه افزایش سرعت یادگیری می‌شود.**

کلید واژه- انتقال یادگیری، مجرد کردن دانش، مفهوم ، یادگیری تقویتی

## ۱- مقدمه

موجودات هوشمند با داشتن هوش طبیعی در تعامل با محیط یاد می‌گیرند که در برابر هر شرایطی از محیط چگونه رفتار کنند. در شاخه‌های مختلف هوش مصنوعی نیز سعی شده است که با الهام گرفتن از نحوه‌ی تعامل سیستم‌های هوشمند طبیعی با محیط روش‌هایی برای توسعه سیستم-های هوشمند مصنوعی ارائه شود. یادگری تقویتی یکی از اما هنوز فاصله قابل توجهی بین قابلیت یادگیری در سیستم‌های هوشمند و سیستم‌های بیولوژیکی دیده می‌شود. یکی از دلایل این مساله استفاده نکردن از دانش بدست آمده در طول یادگیری در یک سیستم هوشمند می-باشد. به این منظور، اخیرا محققان توجه خود را به انتقال دانش در سیستم‌های هوشمند معطوف کرده‌اند.

هدف از انتقال دانش ، یادگیری سریع کار( هدف) بعد از یادگیری کار (مرجع) متفاوت اما مشابه (وابسته) به آن می‌باشد[۱]. انتقال دانش به عامل اجازه می‌دهد که ابتدا یک کار مرجع ساده اولیه را یاد بگیرد و سپس با توجه به آن کار پیچیده‌تری را یاد بگیرد. در صورتی انتقال یادگیری موفق خواهد بود که بعد از استفاده از دانش یادگرفته شده از کار مرجع، یادگیری در کار هدف سریعتر و با عملکرد بهتری صورت گیرد.

اکثر روش‌هایی که تا کنون در راستای انتقال دانش ارائه شده‌اند از نوعی نگاشت بین کارهای مرجع و هدف استفاده کرده‌اند[2]. یافتن چنینی نگاشتی بین کارها ملزم داشتن اطلاعات زیادی در مورد کارها و همچنین وجود و شناسایی تشابهات بین کارها می‌باشد که این امر باعث ایجاد محدودیت در انتقال دانش می‌شود و کار را مشکل می‌سازد. برای حل این مشکل می‌توان از نگاشت در سطح بالایی از تجرید بین دو مساله استفاده کرد. در حالت کلی می‌توان گفت که مقوله‌های انتقال یادگیری و انتزاعی کردن حالت خیلی به هم وابسته و نزدیک بهم هستند[۳]. در انتقال یادگیری عامل می‌کوشد که از دانش کار مرجع در کار هدف استفاده نماید. رسیدن به این هدف نیازمند این است که یک صورت انتزاعی از فضای حالت داشته باشیم تا بتوان دانش کار مرجع را در کارهای دیگر که حوزه‌های کاری متفاوت دارند، استفاده کرد. بنابراین مساله تصمیم‌گیری در مورد انتخاب دانش برای انتقال در حوزه‌های متفاوت

می‌تواند به مساله انتزاعی کردن حالت برای یک مجموعه از حوزه‌های مرجع تبدیل شود. در این مقاله سعی بر این است تا با استفاده از مفاهیم سلسله مراتبی، عامل دانش خود را برای استفاده مجدد در کار دیگر ( معمولا پیچیده‌تر ) مجرد کند. رویکرد ما برای تعریف مفهوم و مدل کردن آن در ذهن عامل یک رویکرد کارکردی است[۴]. یعنی مفاهیم در فضای Q-value ها شکل می‌گیرند. نتایج بدست آمده نشان می‌دهند که روش پیشنهادی برای انتقال دانش نه تنها باعث سرعت در یادگیری می‌شود بلکه حتی با تغییر سنسورهای ربات ( فضای حسی) عامل می‌تواند از دانش قبلی به علت قابلیت تعمیم آن، استفاده کند و بوسیله آن سرعت یادگیری خود را افزایش دهد.

ترتیب مطالب این مقاله بدین صورت می‌باشد: در قسمت دوم مروری خلاصه بر یادگیری تقویتی داریم. قسمت سوم انتقال دانش و راهکارهای موجود و هدف از آن را بررسی می‌کند. بعد از آن در قسمت چهارم تعریفی از مجرد کردن دانش خواهیم داشت. سپس درقسمت پنجم به نحوه ایجاد مفاهیم و یادگیری آنها می‌پردازیم. در قسمت ششم به چگونگی انتقال یادگیری در این مقاله می‌پردازیم و در آخرین بخش نتایج حاصل از شبیه‌سازی آورده شده است.

## ۲- یادگیری تقویتی

در چارچوب یادگیری تقویتی، عامل یادگیرنده در طی عمر خود $t = 0, 1, 2, 3, \ldots$ در تقابل با محیط می‌باشد. در هر مرحله زمانی $t$ عامل حالت جاری محیط $S_t$ را مشاهده می‌کند و بر اساس آن عمل $a$ را انجام می‌دهد که باعث می‌شود محیط به حالت بعدی $s_{t+1}$ منتقل شود و در نتیجه آن عامل پاداش $r$ را از محیط دریافت کند[۵]. در یک سیستم مارکف حالت بعدی محیط و پاداش دریافتی تنها به عمل و حالت قبلی عامل در محیط بستگی دارد. هدف عامل در یادگیری ماکزیمم کردن پاداش بدست آمده در طول زمان می‌باشد. عامل با یادگرفتن نگاشت حالات به اعمال که سیاست نامیده می‌شود، این کار را انجام می‌دهد. بعبارتی هدف عامل انتخاب عمل بطوریکه مقدار پاداش مورد انتظار $E\{\sum_i^{\infty} \gamma^i r_{t+i}\}$ را افزایش دهد، می‌باشد، که $\gamma$ ضریب فراموشی است. یک راه حل مرسوم، تخمین زدن تابع حالت و عمل بهینه یا همان تابع Q که حالات و اعمال را به ماکزیمم پاداش مورد انتظار که با شروع از حالت s و عمل a نگاشت می‌دهد، می‌باشد.

## ۳- انتقال یادگیری

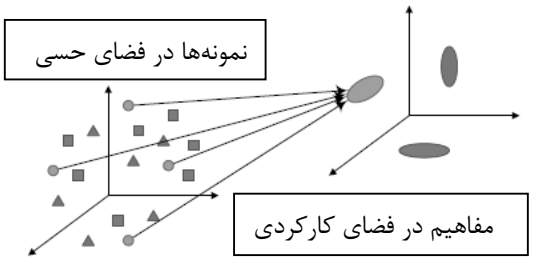
انتقال یادگیری به معنی بکار بردن دانش یادگرفته شده در یک کار ( کار مرجع) به منظور بهبود یادگیری در کار دیگر (کار هدف) می‌باشد. بشر بطور قابل ملاحظه‌ای از دانش‌های یادگرفته شده در کارهای گذشته‌اش برای یادگیری بهتر و سریعتر در کارهای خود بهره می‌برد. بیشتر روش‌هایی که تا کنون در راستای انتقال دانش ارائه شده‌اند از نوعی نگاشت بین کارهای مرجع و هدف استفاده کرده‌اند. یافتن چنینی نگاشتی بین کارها ملزم داشتن اطلاعات زیادی در مورد کارها و همچنین وجود و شناسایی تشابهات بین کارها می‌باشد که این امر باعث ایجاد محدودیت در انتقال دانش می‌شود و کار را مشکل می‌سازد. برای حل این مشکل می‌توان از نگاشت در سطح بالایی از تجرید بین دو مساله استفاده کرد. در این مقاله برای نیل به این هدف از مجرد کردن دانش از طریق یادگیری مفاهیم استفاده شده است.

## ۴- مجردکردن دانش

محیط واقعی که ما انسانها در آن زندگی می‌کنیم سرشار از اطلاعات پیوسته و گسسته که با نا یقینی و نویز همراه است می‌باشد. یادگیری در چنین محیط پیچیده‌ای، بدون استفاده از مکانیزم های خاصی غیر ممکن می‌نماید. یکی از مهم‌ترین حربه هایی که انسان برای برخورد با چنین محیط هایی بکار می‌برد مجرد کردن می‌باشد. مجردسازی مکانیزمی‌است که طی آن فضای ادراکی پیچیده عامل هوشمند به یک فضای ساده‌تر که توسط عامل قابل مدیریت کردن می‌باشد نگاشته می‌شود و از آنجا که سعی می‌شود در این فرآیند تا حد امکان محتوای اطلاعاتی فضای ادراکی اصلی حفظ شود، فضای ادراکی حاصل به نوعی، باز نمایی سطح بالاتر و یا مجردتر از فضای اصلی خواهد بود[۶].

## ۵- یادگیری مفاهیم

یکی از روشهای موجود برای مجردسازی، تقسیم فضای ادارکی عامل به یک سری کلاس های شباهت است بطوری-که هر کلاس حالت‌های مشابه در فضای ادراکی را در خود جای دهد. به هر یک از این کلاس‌های شباهت مفهوم گفته می‌شود. انسانها در تعامل با محیط یادگرفته‌اند که برای درک محیط پیرامونشان تنها به ویژگی های حسی اتکا نداشته باشند و با عمل هایشان محیط را بشناسند. آنها چون قابلیت استخراج مفاهیم مجرد از محیط را دارند، می-توانند یادگیری در یک محیط را به محیط‌های جدید تعمیم دهند. مفاهیمی‌که با اتکا به ارزش اعمال بدست می‌آیند را مفاهیم کارکردی می‌گویند. در این مقاله نشان داده شده است که مفاهیم کارکردی ابزار مناسبی برای مجرد کردن دانش ربات بشمار می‌رود که می‌تواند قابلیت مناسبی برای به اشتراک گذاشتن دانش بین عامل‌هایی که فضای حسی متفاوتی دارند یا بین دو مساله که فضای حسی متفاوتی دارند محسوب شود. شکل (۱) شماتیکی از نمونه ها در فضای حسی و مفاهیم استخراجی از آنها در فضای کارکردی نشان می‌دهد.




شکل ۱: شماتیکی از نمونه ها در فضای حسی و مفاهیم استخراجی از آنها در فضای کارکردی[۴]

## ۶- مجرد کردن دانش و انتقال آن با استفاده از یادگیری تقویتی

مجرد کردن و قابلیت تعمیم دو ویژگی مهم برای سیستم های هوشمند می‌باشد که سرعت و کیفیت یادگیری را بهبود می‌بخشد. در این مقاله برای مجرد کردن دانش، دانش بدست آمده از یادگیری عامل در کار مرجع را بصورت مفاهیم مجرد در می‌آوریم. برای بدست آوردن مفاهیم، بردارهای $Q$ بدست آمده عامل پس از یادگیری در مساله مرجع توسط روش خوشه بندی k-means بصورت دسته هایی مجزا در می‌آیند مراکز این دسته‌ها همان مفاهیم انتقالی هستند که عامل در مساله هدف از آنها برای یادگیری سریعتر استفاده می‌کند. پس از محاسبه مفاهیم آنها را به مساله هدف منتقل می‌کنیم. الگوریتم یادگیری در هر دو مساله مرجع و هدف یادگیری $Q$ می‌باشد. عامل در کار هدف ابتدا یادگیری عادی خود را شروع می‌کند. سپس پس از اینکه به دانش لازم برای تشخیص صحیح مفهوم دست یافت از مفاهیم انتقالی برای تصمیم‌گیری استفاده می‌کند. معیاری که برای این کار در نظر گرفته شده است تعداد تکرار قرار گرفتن عامل در آن حالت می‌باشد. اگر عامل $N_c$ بار یک حالت را مشاهده کرده باشد مجاز به استفاده از مفاهیم برای انتخاب عمل مناسب می‌باشد که این مقدار با سعی و خطا بدست می‌آید. این کار را چون عامل در ابتدای یادگیری دید کافی به مساله ندارد و نمی-تواند تشخیص درستی از مفهومی‌که درآن قرار دارد داشته باشد، انجام می‌دهیم. عامل پس از اینکه توانایی استفاده از مفاهیم را بدست آورد با توجه به بردار ارزش $Q$ که در آن قرار دارد شبیه ترین مفهوم را که نزدیکترین مفهوم به بردار $Q$ می‌باشد را انتخاب می‌کند. فاصله اقلیدسی در اینجا بعنوان شاخص شباهت در نظر گرفته شده است:

$$d_i = \sqrt{(\vec{Q}_s - \vec{Q}_{c_i})^T (\vec{Q}_s - \vec{Q}_{c_i})} \qquad (۱)$$

$$Q_c = \arg\min_{Q_c}(d) \qquad (۲)$$

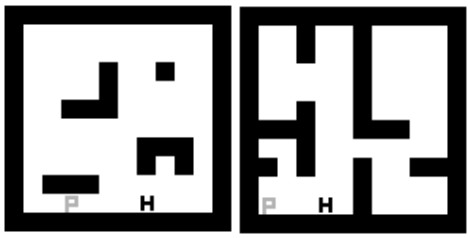
اگر فاصله بردار $Q$ تا نزدیکترین مفهوم از محدوده تعیین شده کم‌تر باشد. عامل با توجه به آن مفهوم تصمیم‌گیری می‌کند در غیر اینصورت با استفاده از بردار $Q$ خود و الگوریتم یادگیری $Q$ عمل مناسب را انتخاب می‌کند. با این کار عامل تنها در حالاتی که دانش مرجع برایش مفید است از دانش انتقالی استفاده می‌کند. با انجام عمل و انتقال محیط به حالت بعدی و پاداشی که محیط به عامل می‌دهد

عامل با استفاده از یادگیری Q مقدار بردار Q خود را با استفاده از رابطه (۳) بروز می‌کند.

$$Q(s,a) \leftarrow (1-\alpha)Q(s,a) + \alpha(r + \gamma \max_{a'} Q(s',a')) \qquad (۳)$$

## ۷- نتایج شبیه‌سازی

مساله ای که برای شبیه‌سازی در نظر گرفته شده، مساله شکار و شکارچی می‌باشد. مساله شکار و شکارچی یکی از مسائل کلاسیک برای مطالعه و مقایسه روش‌های متفاوت یادگیری در هوش مصنوعی می‌باشد[۷].



شکل ۲: نمایی از دو محیط شبیه‌سازی ، شکل سمت چپ محیط شبیه‌سازی برای مساله مرجع و سمت راست برای مساله هدف

شکار و شکارچی در یک محیط دو بعدی مربعی شکل که هر بعد آن به ۱۲ قسمت تقسیم شده است، واقع شده‌اند. دورتا دور محیط دیوار قرار گرفته است و تعدادی مانع در آن قرار داده شده است. عامل شکارچی در هر قدم یکی از پنج عمل بالا، پائین، راست، چپ و یا تیر زدن را با احتمالی که با تابع بولتزمن مشخص می‌شود انتخاب می‌کند. شکارچی بایستی بدون برخورد به مانع و در کمترین زمان ممکن شکار را بگیرد یا آن را با تیر بزند. برای گرفتن شکار کافی است که شکارچی در یکی از خانه های مجاور شکار قرار بگیرد. الگوریتم یادگیری شکارچی یادگیری Q می‌باشد و شکار هم با استفاده از روش میدان پتانسیل از شکارچی فرار می‌کند. در هر گام یادگیری پس از اینکه عامل یادگیرنده ( شکارچی ) عمل حاصل از تصمیم گیری خود را انجام داد بر اساس اینکه به چه ناحیه ای از محیط منتقل شده است از محیط پاداش دریافت می‌کند که این تابع پاداش بدین صورت می‌باشد:

بازای هر حرکت بدون نتیجه پاداش ۱- به شکارچی داده می‌شود. اینکار باعث می‌شود که شکارچی با کمترین حرکت شکار را متوقف کند. در صورتیکه شکارچی به مانع برخورد کند پاداش ۴- به او داده می‌شود. هزینه‌ای که شکارچی بازای تیرهای بدون نتیجه می‌پردازد ۶- می‌باشد. این هزینه باعث می‌شود که شکارچی فقط در صورت نیاز ( در تیر رس بودن شکار ) تیراندازی کند.۲۰+ امتیاز پاداش را شکارچی وقتی که شکار را بدون تیر زدن متوقف می‌کند ( در یکی از خانه های مجاور شکار قرار می‌گیرد ) دریافت می‌کند. اگر شکارچی با تیر شکار را متوقف کند پاداش ۳۰+ می‌گیرد.

هر دوره یادگیری با متوقف شدن شکار پایان می‌یابد. پس از آن دوباره شکار و شکارچی بطور تصادفی در محیط قرار داده می‌شوند و دوره‌ای جدیدی آغاز می‌شود. این کار تا تمام شدن تعداد تکرارهای در نظرگرفته شده برای یادگیری ادامه می‌یابد. پارامترهای یادگیری بصورت زیر تنظیم می‌شوند:

نرخ یادگیری($\alpha$) بطور نزولی با افزایش تعداد تکرارها از مقدار ۰.۸ تا ۰.۲ کاهش می‌یابد.

ضریب فراموشی ( $\gamma$ ) برابر با ۰.۹ در نظر گرفته شد.

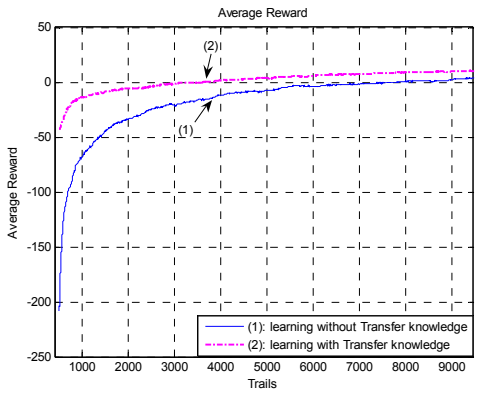
همانطور که قبلا گفته شد دانش بدست آمده عامل پس از یادگیری در کار مرجع ابتدا بصورت مجرد در آمده و سپس به کار هدف منتقل می‌شود. برای مجرد کردن دانش و در آوردن مفاهیم بردارهای ارزش Q نرمالیزه شده را با استفاده از روش خوشه بندی k-means دسته‌بندی می‌کنیم. مراکز خوشه‌ها که مفاهیم انتقالی هستند در این مساله به صورت زیر می‌باشند:

| concepts | Value of 1th action | Value of 2th action(go right) | Value of 3th action(go left) | Value of 4th action(go down) | Value of fifth action(shot) |
|---|---|---|---|---|---|
| $C_1$ | 0.064 | 0.069 | 0.063 | 0.989 | 0.0156 |
| $C_2$ | 0.079 | 0.076 | 0.083 | 0.076 | 0.9945 |
| $C_3$ | 0.067 | 0.069 | 0.978 | 0.062 | 0.0145 |
| $C_4$ | 0.987 | 0.068 | 0.077 | 0.061 | 0.0173 |
| $C_5$ | 0.071 | 0.965 | 0.066 | 0.064 | 0.0168 |
| $C_6$ | 0.750 | 0.748 | 0.755 | 0.742 | 0.0802 |

همانطور که ملاحظه می‌شود ۶ مفهوم از محیط بدست آمده است. با توجه به مراکز بدست آمده، در هر بعد تنها یکی از

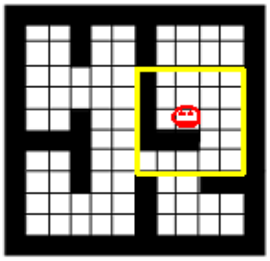
اعمال که زیر آنها خط کشیده شده است از احتمال بالایی برای انتخاب برخوردارند.

با توجه به بردارهای مفاهیم بدست آمده می‌بینیم که هر مفهوم یک عمل خاص را توصیه می‌کند مثلا مفهوم اول عمل چهارم که در اینجا پایین آمدن است را توصیه می‌کند مفهوم دوم که نشانگر در تیررس بودن شکار می‌باشد عمل تیر زدن را توصیه می‌کند. عامل در مساله هدف پس از طی یک دوره یادگیری، در حالت‌هایی که بیش از ۵ بار در آنها قرار گرفته از دانش منتقل شده برای تصمیم‌گیری کمک می‌گیرد. نتایج حاصل از شبیه‌سازی در شکل (۳) آمده است. لازم به توضیح است که نمودارها، حاصل متوسط-گیری از ۵ بار اجرا هستند تا اطلاعات بدست آمده قابل اعتمادتر شوند و اثر حالت‌های خاص کم‌تر شوند.



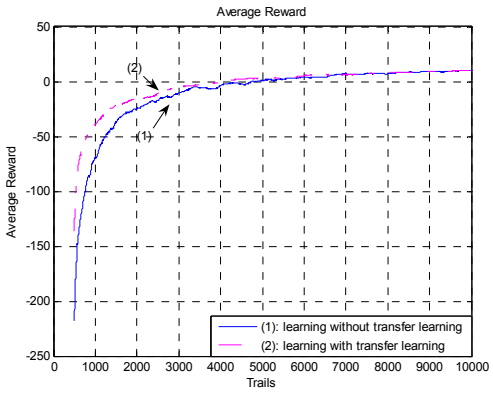

شکل ۳- مقایسه نمودارهای پاداش متوسط: (۱) یادگیری بدون انتقال دانش و (۲) یادگیری با انتقال دانش

برای نشان دادن قابلیت تعمیم مفاهیم انتقالی و اینکه دانش انتقالی به فضای حسی عامل بستگی ندارد، سنسورهای ربات را در کار هدف عوض کرده و از همان دانشی که ربات با استفاده از سنسورهای قبلی خود بدست آورده به یادگیری در کار هدف می‌پردازیم. سنسورهای اولیه ربات موقعیت x و y ربات شکارچی و x و y شکار را می‌دهند. در صورتیکه سنسورهای ثانویه ربات همانطورکه در شکل (۴) مشاهده می‌شود، تنها موقعیت خانه‌های اطراف ربات تا شعاع دو خانه مجاور را به ربات می‌دهد.



شکل ۴: ربات با سنسورهای جدید در محیط یادگیری

اما با وجود اینکه سنسورهای دو کار تغییر کرده باز هم مفاهیم بدست آمده از کار مرجع یادگیری ربات را بهبود بخشیده است. شکل(۵) نمودار پاداش متوسط را نشان می-دهد.




شکل۳- مقایسه نمودارهای پاداش متوسط: (۱) یادگیری بدون انتقال دانش و (۲)یادگیری با انتقال دانش در دو مساله با فضای حسی متفاوت

در آخرین قسمت از شبیه‌سازی‌های انجام شده انتقال یادگیری بین یک محیط مارکف (MDP) و یک محیط مشاهده پذیر جزیی مارکف (POMDP) بررسی می‌شود. در یک محیط POMDP بعضی از حالت‌های متفاوت محیط از نظر عامل یکی می‌شوند و این امر باعث می‌شود که عامل به یادگیری مناسبی دست پیدا نکند. در این جا ما با انتقال دانش کسب‌شده از کار مرجع که MDP بوده باعث بهبود یادگیری عامل در محیطی که برای عامل POMDP است، می‌شویم. شکل (۶) حاکی از این امر می‌باشد.

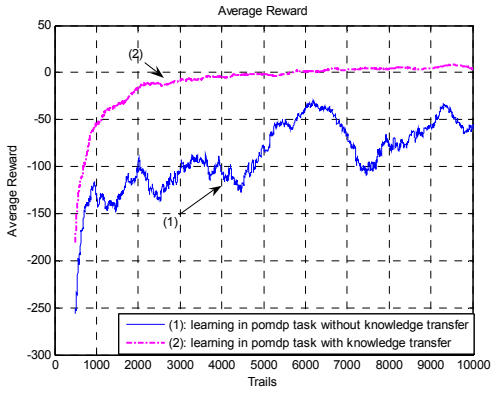
[3] Thomas J.Wash and Michael L. Litman. Transfering state Abstraction between MDPs. In ICML-07 conference.

[4] Hamide Vosoughpour, Majid Nili Ahmadabadi, Maryam S. Mirian, Babak Nadjar Araabi. Hierarchical Functional Concept Formation using Reinforcement Learning. In ICTA2007 conference.

[5] R. S. Sutton and A. G. Barto. Reinforcement Learning: An Introduction. MIT Press, Cambridge, MA, 1998.

[6]. Hadi Firouzi, Majid Nili Ahmadabadi, Babak N. Araabi. A Probabilistic Reinforcement-Based Approach to Conceptualization. In International Journal of Intelligent Technology (IJIT), Volume 3, pp 48-55, 2008

[7] M. Tan, "Multi-agent reinforcement learning: Independent vs. cooperative agents," in Proc. Tenth Int. Conf. Machine Learning, Amherst, MA, June 1993.




شکل ۶- مقایسه نمودارهای پاداش متوسط: (۱) یادگیری بدون انتقال دانش و (۲) یادگیری با انتقال دانش بین دو مساله MDP و POMDP

## ۸- نتیجه‌گیری

دراین مقاله به بررسی انتقال دانش بین دو کار در حالت-های مختلف پرداخته شد و نشان داده شد که اگر دانش بدست آمده از کار اول را بصورت مجرد در آوریم براحتی می‌توانیم آن را در یک کار مشابه دیگر استفاده کنیم حتی اگر فضای حسی دو کار با یکدیگر متفاوت باشد. در این روش دیگر نیازی به نگاشت یک به یک بین حالت و عمل عامل در دو کار نیست و مشکل پیدا کردن شباهت بین کارها و نگاشت بین آنها که در دیگر روشهای انتقال دانش وجود دارد در اینجا وجود ندارد. همچنین روش پیشنهادی باعث بهبود پاداش بدست آمده و افزایش سرعت یادگیری می‌شود.

## مراجع

[1] Matthew E. Taylor and Peter Stone. Behavior Transfer for value function Based Reinforcement Learning. In conference on Autonomous agents and Multi agent System (AAMAS-05) pp. 53-59

[2] Yaxin Liu and Peter Stone. Value-Function- Based Transfer for Reinforcement Learning Using Structure Mapping. in Proceedings of the Twenty-First National Conference on Artificial Intelligence (AAAI-06),