

بسمه تعالی



دانشگاه تهران

پردیس دانشکده هایفنی

دانشکده مهندسی برق و کامپیوتر

پیشنهاد و فرم حمایت از پایان نامه تحصیلات تکمیلی

☐

دکتری

☒

کارشناسی ارشد

شماره مرجع : *

*شماره مرجع، توسط معاونت پژوهشی پردیس دانشکده‌های فنی هنگام صدور ابلاغ درج خواهد شد.

1- خلاصه اطلاعات پایان نامه

عنوان پایان نامه به زبان فارسی:

بررسی تعامل دوسیستم یادگیری دارای مدل و بدون مدل

عنوان پایان نامه به زبان انگلیسی:

نوع پایان نامه: بنیادی ☒ کاربردی ☐ توسعه‌ای ☐

بردرس/دانشکده: فنی دانشکده/گروه: مهندسی برق و کامپیوتر

مقطع تحصیلی: کارشناسی ارشد رشته و گرایش تحصیلی: هوش مصنوعی و رباتیک

تاریخ پیشنهاد: تاریخ تصویب:

2- اطلاعات اساتید راهنما و مشاورین

نوع مسئولیت	نام و نام خانوادگی	مرتبه علمی	محل خدمت	امضاء
استاد راهنما (مجری)	دکتر نیلی	استاد	دانشگاه تهران - دانشکده فنی	
استاد راهنمای دوم (حسب نیاز)				
استاد مشاور				
استاد مشاور دوم (برای دکتری)				

3- اطلاعات دانشجو

نام و نام خانوادگی: یاسمن رازقی شماره دانشجویی: 810194508 رشته و گرایش

تحصیلی: هوش مصنوعی و رباتیک دانشکده: مهندسی برق و کامپیوتر مقطع تحصیلی: کارشناسی ارشد

پست الکترونیک: yasamanrazeghi7@gmail.com تلفن ثابت: 02122053854 تلفن همراه: 09124331187

4- مشخصات موضوعی پایان نامه

تعریف مسأله، هدف و ضرورت اجرا (حداکثر سه صفحه)

یادگیری تقویتی

یادگیری تقویتی مدلی است که در سالیان اخیر برای بررسی یادگیری موجودات زنده به کار گرفته شده است و با توجه به شواهد و نتایج بدست آمده موفقیت چشم‌گیری در این حوزه داشته است. این مدل توانسته است بسیاری از رفتارهای انسان و دیگر موجودات زنده را مانند ایجاد عادات رفتاری توجیه کرده و پاسخ مدلسازی قابل قبولی برای این رفتارها ارائه کند. از طرفی دیگر شواهد و اطلاعاتی که از مطالعات آسیب‌های مغزی، دستکاری‌های دارویی و ثبت‌های متفاوت مغزی به دست آمده‌اند می‌توانند این چهارچوب را تأیید کنند. همچنین این مدل‌ها توانسته است سیستم‌های یادگیری مغز را که بر مبنای دوپامین کار می‌کنند و یکی از پایه‌ای‌ترین نقش‌ها در فرایند یادگیری را دارد توجیه کند [1][2][3]. این موفقیت در توجیه شواهد یادگیری در انسان موجب تقویت این حوزه و جذب بسیاری به این حوزه شده است.

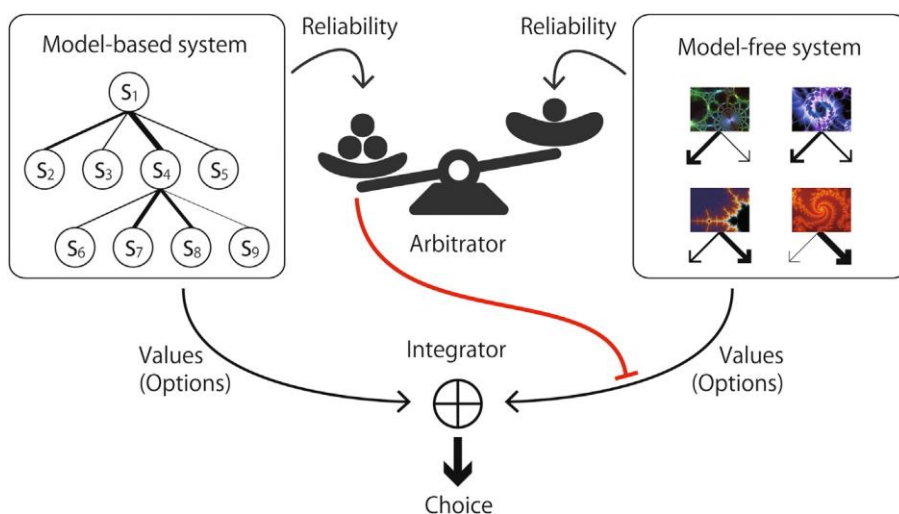
سیستم‌های یادگیری

نظریات متفاوتی در رابطه با سیستم‌های متفاوت یادگیری در انسان ارائه شده است. بر طبق شواهد به نظر می‌رسد یکی از این تئوری‌ها تطابق بیشتری با داده‌های رفتاری و مغزی انسانی داشته و در حال حاضر تعداد بسیاری از دانشمندان این حوزه را به خود درگیر کرده است. براساس این تحقیقات به نظر می‌رسد که پستانداران از دو نوع سیستم برای یادگیری انتخاب درست استفاده می‌کنند: سیستم دارای مدل و سیستم بدون مدل. در سیستم اول که سیستم دارای مدل شناخته می‌شود عامل تلاش می‌کند که بر مبنای اطلاعاتی که از محیط پیرامون خود به دست آورده است، محیط پیرامون را در غالب یک مدل ذهنی یاد بگیرد و با استفاده از این مدل ذهنی یک درخت تصمیم‌گیری تشکیل دهد و با محاسبه امید ریاضی مسیرهای متفاوت در این درخت بهترین مسیر را انتخاب کند. این روش برای رسیدن به هدف از نظر آماری بهینه است اما حجم محاسبات و در نتیجه تلاش ذهنی زیادی نیاز دارد. در سیستم دوم که سیستم بدون مدل است، عامل مدلی از محیط ندارد و با بازخورد از محیط انتخاب‌های خود را بهینه می‌کند. به این صورت که عامل ابتدا برای هر یک از انتخاب‌های خود بسته به شرایط یک ارزش ذهنی در نظر دارد، و بر اساس آن ارزش ذهنی اولیه انتخاب خود را انجام داده و از محیط پاداشی دریافت می‌کند و از این پاداش یا بازخورد برای بهبود ارزش اعمال استفاده می‌کند. این پاداش می‌تواند با تصور پیشین عامل متفاوت باشد که ازین تناقض برای بهبود ارزش‌های ذهنی خود استفاده می‌کند تا اگر عامل در شرایط محیطی مشابه قرار گرفت تخمین‌های واقع‌بینانه‌تری از ارزش انتخاب‌های خود داشته باشد. این روش، روش بهینه‌ای نیست اما هزینه‌ی محاسباتی پایین‌تری

دارد و بسته به محیط می‌تواند پاداش‌هایی به اندازه‌ی کافی خوب برای عامل داشته باشد. همچنین هزینه‌ی محاسباتی پایین این سیستم سرعت تصمیم‌گیری را نیز می‌تواند بسیار بالا برد. حضور این دو سیستم در کنار هم باعث بروز انعطاف‌پذیری در یادگیری در شرایط یادگیری متفاوت و در محیط‌های گسترده‌ی متفاوت شود. شواهد مغزی بسیاری هم برای کد شدن نتایج این دو سیستم در مغز و نحوه‌ی تعامل و همکاری آن‌ها وجود دارد [4].

تعامل این دو سیستم با تعاریف متفاوتی بیان شده‌است، در یکی از این تعاریف دو سیستم هم‌زمان رقابت می‌کنند و یک سیستم میانجی با توجه به اطمینان‌پذیری این دو سیستم از بین آنها انتخاب می‌کند. با این معنا که در لحظه هر سیستمی که با توجه به تصمیم‌های پیشین خود اطمینان‌پذیری بیش‌تری داشته باشد برای تصمیم‌گیری استفاده می‌شود. شماتیک این تعامل را می‌توان در شکل شماره‌ی یک دید [17].

در تعریفی دیگر این دو سیستم با هم همکاری می‌کنند. به این شکل که سیستم بدون مدل به‌صورت آنلاین سبب بروز جنبه‌های مختلف رفتار انسانی و سیستم دارای مدل به‌صورت آفلاین همواره در حال تصحیح رفتار سیستم بدون مدل است [18][19].



شکل شماره ی یک یکی از نظرهای تعامل دو سیستم را نشان می دهد

مطالعات در مورد نحوه‌ی کار این دو سیستم و بروز نشانه‌های مغزی و رفتاری این دو و همچنین تعامل رقابتی و یا همکاری این دو سیستم در موجودات زنده همچنان ادامه دارد. که این پژوهش هم قصد دارد تا در این راستا گام بردارد.

هدف این پژوهش ارائه‌ی یک تعریف و پنجره‌ی جدید از حضور یک سیستم مرتبه‌ی بالاتر هنگام تعامل دو سیستم است. این سیستم که می‌تواند خود نوعی سیستم دارای مدل تعبیر شود، با روش کنترلی تطبیقی می‌تواند به تنظیم تعامل رفتار دارای مدل و بدون مدل بپردازد. ارائه‌ی این مدل در قالب یک سیستم مرتبه بالای یادگیرنده می‌تواند آغازی برای توجیه رفتارهای پیچیده‌ی انسانی باشد.

بیان سوال اصلی

سوال اصلی این پروژه را می‌توان به بررسی تعامل بین این دو سیستم یادگیری بر مبنای مدل و بدون مدل نسبت داد و اینکه آیا این تعامل ناشی از یک رفتار تطبیقی کنترلی سطح بالاتر در انسان بروز می‌کند یا خیر. تسک‌های بسیاری برای بررسی این دو سیستم و ارائه‌ی پارامتری برای جداسازی و تعامل این دو سیستم ارائه شده است اما هیچ تحقیقی به بررسی حضور یک سیستم سطح بالاتر که باعث بروز رفتار تطبیقی کنترلی در تعامل این دو سیستم می‌شود، نپرداخته است. در این پروژه می‌خواهیم این روند تعامل بین دو سیستم را بررسی کنیم و مدل بهتری از آنچه تا به حال به توضیح این تعامل پرداخته دهیم.

ضرورت اجرای این پژوهش:

این پژوهش از چندین جهت می‌تواند جالب باشد اول اینکه شناخت بهتر و درست‌تر ما نسبت به تعامل سیستم‌های یادگیرنده‌ی انسانی می‌تواند ما را به سمت استفاده‌ی بهتر از این سیستم‌ها در راستای یادگیری بهتر و طراحی دوره‌های یادگیری مفیدتر برای رسیدن به رفتارهای منطقی سوق دهد. همچنین در مواردی که به دلایلی مثل بروز اعتیاد، بیماری‌های متفاوت و عوامل مداخله‌گر دیگر رفتار یکی از این سیستم‌ها دچار اختلال شود با دانستن نحوه‌ی تعامل می‌تواند نحوه‌ی یادگیری عامل را به سمت استفاده از سیستم مداخله‌شده و رسیدن به انتخاب‌های عاقلانه‌تر سوق داد. همچنین نگاه کنترلی تطبیقی به تعامل این دو سیستم یک نگاه کاملاً جدید بوده که پنجره‌ی جدیدی از رفتار دو سیستم را باز خواهد کرد که هم زمان می‌تواند شواهد دال بر رقابت دو سیستم و شواهد دال بر همکاری دو سیستم را توضیح دهد.

چشم‌اندازهایی که برای انجام این پژوهش در نظر گرفته شده‌است شامل مطالعه‌ی تسک‌های موجود و استفاده از آنها برای طراحی یک تسک مارکوف جداکننده‌ی دو سیستم، جمع‌آوری داده از افراد، مدل‌سازی تعامل دو سیستم با توجه به پارامترهای فرایادگیری، ارائه‌ی شواهد مغزی برای مدل‌سازی است.

۱- طراحی یک تسک مارکوف جداکننده‌ی دو سیستم یادگیرنده
برای ارائه‌ی روند تغییر دو سیستم و تعامل آن‌ها نیاز به تسکی که بتواند عامل را به سمت استفاده از هر یک از این دو سیستم متمایل کند. همچنین نیاز به محیطی داریم که تمایز رفتاری بین این دو سیستم قابل مشاهده باشد. برای این کار روش‌های متفاوتی ارائه شده‌است که یکی از موفق‌ترین آن‌ها در [17][18] است. در این قسمت باید با استفاده از روش‌های موجود و همچنین محدودیت‌ها مدل‌سازی مسئله به طراحی آزمایش مورد نیاز خود برای دنبال کردن الگوی تعامل بین دو سیستم بپردازیم.

۲- جمع‌آوری داده
بعد از مرحله‌ی طراحی تسک در طول آزمایش اطلاعات رفتاری افراد ثبت شده تا در مرحله‌ی تحلیل مورد بررسی قرار گیرد.

۳- مدل‌سازی تعامل دو سیستم
مدل‌سازی قدم مهمی در انجام این پروژه است، این مدل‌سازی با توجه به پارامترهای فرایادگیری نسبت به ارائه‌ی مدلی برای تعامل بین دو سیستم یادگیرنده را با توجه به ادبیات یادگیری تقویتی انجام می‌دهد.

۴- تصویربرداری عملکردی تشدید مغناطیسی
در این مرحله تلاش خواهیم کرد تا شواهد مغزی‌ای برای حضور یک سیستم تعاملی مرتبه‌ی بالاتر برای تعامل این دو سیستم با گرفتن داده‌های تصویربرداری تشدید مغناطیسی در حین انجام تسک، ارائه دهیم.

سیستم‌های یادگیری

یکی از نتایج مهم تحقیقات انجام گرفته در چهارچوب یادگیری تقویتی حضور چندین روش یادگیری در عوامل یادگیری است. بسیاری از حضور دو سیستم یادگیرنده با نام‌های مختلف توضیح می‌دهند. بر اساس این چهارچوب «دو سیستم یادگیری» بسیاری از ابعاد متفاوت رفتار انسانی قابل توجیه می‌شود که توضیح یکی از موردقبول‌ترین این تئوری‌ها در قسمت مطالعات پیشین ارائه شده است. مواردی

مثل نحوه‌ی ایجاد عادات و یا انواع بایاس‌های رفتاری توسط این دو سیستم قابلیت پشتیبانی دارد. یکی از تعاریف این دو سیستم، سیستم یادگیری بر مبنای مدل و بدون مدل هستند. پژوهش‌ها در چند سال اخیر به سمت نحوه‌ی کار این دو سیستم و شناسایی روش‌هایی برای تمایز این دو سیستم ارائه شده است. در این میان تسک‌هایی متفاوتی که با تحلیل داده‌های رفتاری و مغزی به بررسی و مدل‌سازی و یافتن شواهد برای نحوه‌ی بروز کار این دو سیستم در عامل طراحی شده‌اند. [13][14][15]

فرایادگیری

در حوزه‌ی یادگیری ماشین فرایادگیری به معنای یادگیریِ یادگیری است. به صورت شهودی الگوریتم‌های فرایادگیری از تجربه‌ها استفاده می‌کنند تا تا جنبه‌های متفاوتی از الگوریتم یادگیری خود را بهبود بخشند. این یادگیری بهبود یافته از الگوریتم یادگیری اولیه عملکرد بهتری خواهد داشت.

این مفهوم اولین بار در سال ۱۹۷۹ بیان شد و به بررسی روند عامل‌های یادگیرنده‌ای که با استفاده از نوعی کنترل روی فرآیند یادگیری خود به بهبود این روند می‌پردازند، پرداخت. [5] می‌توان فرایادگیری را آگاهی نسبت به فرآیند یادگیری در ناخودآگاه عامل مستقل از دانش شخصی عامل تعریف کرد. در واقع این مفهوم می‌تواند به خودکار کردن تصمیم‌های انسان و بهینه کردن این تصمیم‌ها در حین یادگیری بپردازد.

بعضی از فلاسفه اعتقاد دارند که روش‌های علمی در واقع یکی از حالت‌های پیاده‌سازی فرایادگیری هستند. [6][7]

این حوزه همچنین دارای مدل‌سازی‌ها و فرمول‌بندی‌های ریاضی مخصوص به خود است که می‌تواند رده‌ی فرایادگیری، شروع فرآیند فرایادگیری، پروسه‌ی اضافه شدن تجربه، دامنه‌ها، پارامترهای فرایادگیری، دانش‌های قبلی عامل و الگوریتم‌های متفاوت یادگیری را مدل کند. [8][9][10][11]

- [1] Montague, P. Read, Peter Dayan, and Terrence J. Sejnowski. "A framework for mesencephalic dopamine systems based on predictive Hebbian learning." *The Journal of neuroscience* 16.5 (1996): 1936-1947.
- [2] Fiorillo, Christopher D., Philippe N. Tobler, and Wolfram Schultz. "Discrete coding of reward probability and uncertainty by dopamine neurons." *Science* 299.5614 (2003): 1898-1902.
- [3] Daw, Nathaniel D., and Kenji Doya. "The computational neurobiology of learning and reward." *Current opinion in neurobiology* 16.2 (2006): 199-204.
- [4] Dolan, Ray J., and Peter Dayan. "Goals and habits in the brain." *Neuron* 80.2 (2013): 312-325
- [5] Norton, L. & Walters, D (2005). Encouraging meta-learning through personal development planning: first year students' perceptions of what makes a really good student. PRIME (Pedagogical Research In Maximising Education), in-house journal, Liverpool Hope University, 1 (1) 109-124.
- [6] Breiman, L., 1996. Bagging Predictors. *Machine Learning*, 24(2), 123-140.
- [7] Anderson, M.L. & Oates, T., 2007. A review of recent research in metareasoning and metalearning. *AI Magazine*.
- [8] Brazdil, P.B. et al., 2009. *Meta-Learning: Applications to Data Mining*, Springer.
- [9] Cramer, N.L., 1985. A Representation for the Adaptive Generation of Simple Sequential Programs. In J. J. Grefenstette *Proceedings of an International Conference on Genetic Algorithms and Their Applications*. Hillsdale NJ.
- [10] Gagliolo, M. & Schmidhuber, J., 2006. Learning Dynamic Algorithm Portfolios. *Annals of Mathematics and Artificial Intelligence*, 47(3-4), 295-328.
- [11] Giraud-Carrier, C., Vilalta, R. & Brazdil, P., 2004. Introduction to the special issue on meta-learning. *Machine Learning*, 54(3), 187–193.
- [12] Dayan, Peter, and Yael Niv. "Reinforcement learning: the good, the bad and the ugly." *Current opinion in neurobiology* 18.2 (2008): 185-196.
- [13] Dolan, Ray J., and Peter Dayan. "Goals and habits in the brain." *Neuron* 80.2 (2013): 312-325.
- [14] Kahneman, Daniel. *Thinking, fast and slow*. Macmillan (2011).
- [15] Dayan, Peter. "Rationalizable irrationalities of choice." *Topics in cognitive science* 6.2 (2014): 204-228.
- [16] Daw, Nathaniel D., et al. "Model-based influences on humans' choices and striatal prediction errors." *Neuron* 69.6 (2011): 1204-1215.
- [17] Lee, S.W., Shimojo, S., and O'Doherty, J.P. (2014). *Neuron* 81, this issue, 687–699.
- [18] N.D. Daw, Y. Niv, P. Dayan "Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control" *Nat Neurosci*, 8 (2005), pp. 1704–1711
- [19] S.J. Gershman, A.B. Markman, A.R. Otto "Retrospective revaluation in sequential decision making: a tale of two systems" *J Exp Psychol Gen*, 143 (2014), pp. 182–194

5- مصوبه شورای پژوهشی و تحصیلات تکمیلی دانشکده مهندسی برق و کامپیوتر

1-5- فرم پیشنهاد و حمایت از پایان نامه در تاریخ..... در شورای پژوهشی و تحصیلات تکمیلی دانشکده /گروه		
مطرح و نظرشورا به شرح زیر اعلام می شود:		
<input type="checkbox"/> تصویب شد	<input type="checkbox"/> نیاز به اصلاح دارد	<input type="checkbox"/> به تصویب نرسید
2-5- عنوان طرح جامع تحقیقات استاد راهنما:		
سیستم های اطلاعاتی و محیط های هوشمند		
3-5- آیا پایان نامه پیشنهادی مرتبط با طرح جامع تحقیقات استاد راهنما/مشاور/گروه آموزشی / دانشکده می باشد:		
<input type="checkbox"/> خیر	<input type="checkbox"/> بلی	
امضا استاد راهنما		
مضاء رئیس / معاون پژوهشی و تحصیلات تکمیلی دانشکده مهندسی		

شماره:
تاریخ:
معاون محترم آموزشی و تحصیلات تکمیلی پردیس دانشکده های فنی
با سلام و احترام،
فرم پیشنهاد و حمایت از پایان نامه کارشناسی ارشد / رساله دکتری آقای / خانم
با عنوان
به راهنمایی آقای / خانم دکتر
در شورای پژوهشی و تحصیلات تکمیلی دانشکده مهندسی مورخ به تصویب رسید.
خواهشمند است دستور فرمایید اقدامات مقتضی انجام شود.
امضاء رئیس / معاون پژوهشی و تحصیلات تکمیلی دانشکده مهندسی

شماره:

تاریخ:

معاون محترم پژوهشی پردیس دانشکده های فنی

با سلام و احترام ،

به پیوست فرم پیشنهاد و حمایت از پایان نامه تحصیلات تکمیلی با مشخصات مذکور که به تصویب شورای پژوهشی و تحصیلات تکمیلی دانشکده مهندسی رسیده است، جهت دستور اقدام مقتضی تقدیم می شود.

امضاء معاون آموزشی و تحصیلات تکمیلی پردیس دانشکده های فنی

رونوشت: معاون محترم پژوهشی و تحصیلات تکمیلی دانشکده مهندسی : جهت اطلاع و پیگیری