

# بهبود مدل سازی رفتار دانشجو در سیستم شبه آموزشی با استفاده از اتوماتای یادگیر

لیلا جوادی<sup>۱</sup>؛ بهروز معصومی<sup>۲</sup>؛ محمدرضا میبدی<sup>۳</sup>

## چکیده

در دهه اخیر مساله تعیین نحوه یادگیری و کسب دانش، توجه محققان را به خصوص در بخش سامانه های آموزشی هوشمند جلب کرده است. یکی از اجزاء این سامانه ها مدل سازی دانشجو است. مدل دانشجو، نمایشی از رفتار و وضعیت دانشجو است. در هنگام استفاده دانشجو از سیستم، دنباله ای از اعمال انجام شده او ثبت می شود و سیستم تلاش میکند تا بر اساس آن، مدل یادگیری دانشجو را تعیین کند. این مقاله یک دیدگاه جدید برای مدل کردن رفتار دانشجو در یک سیستم شبه آموزشی با استفاده از اتوماتای یادگیر بیان می کند. در مدل پیشنهادی، مدل دانشجو توسط یک اتوماتای سطح بالاتر به نام Meta\_LA که مدل یادگیری دانشجو را مشخص می کند، تعیین می شود. Meta\_LA در حقیقت اتوماتای یادگیر را به عنوان مکانیزم یادگیری برای تعیین اینکه دانشجو یادگیرنده کند، نرمال یا سریع می باشد، بکار می گیرد. در این مقاله با به کارگیری اتوماتای یادگیر Pursuit دقت مدل یادگیری دانشجو افزایش یافته است.

## کلمات کلیدی

اتوماتای یادگیر، مدل سازی دانشجو، سیستم های شبه آموزشی

## Improving Student's Behavior Modeling in a Tutorial-Like System Using Learning Automata

Leila javadi ; behrooz masoumi; Mohammad reza Meybodi

### ABSTRACT

For acquiring knowledge and the way of learning determination, especially in intelligent tutorial systems draw researchers attention in the recent decade. Student modeling is one of the components of these systems. This model shows student's behavior and state. when the student use the system, his (her) actions is recorded and the system try to determine the student learning model according to them. This paper gives a new opinion and viewpoint to model student's behavior in a tutorial-Like system using learning automata .in the proposing model, student model, determined by high level automata called Meta-LA which shows student learning model. Meta-LA actually use learning automata as learning mechanism to show how the student is slow ,normal or quick in term of learning. This paper shows student learning model accuracy increases using Pursuit learning automata.

### KEYWORDS

Learning automata (LAs), student modeling, tutorial-like systems

<sup>۱</sup> دانشجوی کارشناسی ارشد، گرایش نرم افزار، دانشگاه آزاد اسلامی قزوین. Leila\_javadi17@yahoo.com

<sup>۲</sup> استادیار دانشکده مهندسی کامپیوتر و فناوری اطلاعات دانشگاه آزاد اسلامی قزوین. Masoumi@Qiau.ac.ir

<sup>۳</sup> استاد دانشکده مهندسی کامپیوتر و فناوری اطلاعات دانشگاه صنعتی امیرکبیر. mmeybodi@aut.ac.ir

## ۱. مقدمه

در طول بیست سال اخیر حضور گسترده سیستم های رایانه ای در امور آموزشی هرروز بیشتر شده است. بیشتر سیستم های آموزش به کمک کامپیوتر قبلی، مطالب آموزشی را به صورت یکسان و بدون در نظر گرفتن علائق و پیش زمینه علمی و شخصیتی یادگیرنده عرضه می نمایند، در نتیجه مدل ذهنی و توانایی کاربر مورد توجه قرار نگرفته و مراحل و شیوه آموزش، تناسبی با نیاز و علاقه وی ندارد. یک نظریه مهم در تعلیم و تربیت بیان می کند که هر فرد سبک یادگیری مخصوص خود را دارد. بنابراین در یک سیستم آموزشی هوشمند با استفاده از تکنیک های هوش مصنوعی، براساس نیاز یادگیرنده، موضوع آموزشی به صورت اختصاصی برای وی آماده و در بهترین روش ارائه می گردد. سامانه های آموزشی هوشمند، نسل جدیدی از سامانه های آموزشی هستند که شامل تکنیک های هوش مصنوعی و متدهای ارائه دانش بوده و به صورت واکنش پذیری بر اساس نیاز یادگیرنده توصیف شده که با هدف بهبود آموزش و یادگیری انسان ها بوجود آمده اند. این سیستم ها امکان تقلید سازی از یک معلم انسانی را بوجود آورده به طوریکه می توان آنچه را که تدریس می شود و همچنین چگونگی تدریس را کنترل نمود [۲۱]. برای سیستم های آموزشی هوشمند با توجه به تعابیر مختلف از مفهوم هوشمند سازی و نیز حوزه های کاربردی مختلف، معماری ها و اجزاء گوناگونی پیشنهاد شده است. آنچه در تمامی این طرح ها تقریباً مشترک است، مفهومی از تعامل بین یادگیرنده و مدل هایی از دانش دامنه مورد نظر، افراد متخصص و مدل استخراج شده از یادگیرنده می باشد. در حالت کلی این سیستم ها اساساً شامل سه مولفه اصلی هستند و گاهی مولفه واسطه کاربری نیز به آن اضافه می شود [۳ و ۴]. سه مولفه تشکیل دهنده سیستم آموزشی هوشمند شامل مدل دامنه (دانش دامنه)، مدل دانشجو و مدل آموزشی<sup>۱</sup> می باشند. Self این مولفه ها را به عنوان معماری سه جانبه برای سیستم آموزشی هوشمند تعریف کرده است که به صورت: what (مدل دامنه)، who (مدل دانشجو) و how (مدل آموزشی و تدریس) است [۵].

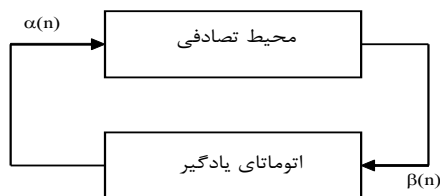
به کارگیری یادگیری ماشین در بهبود سیستم های آموزشی در تعدادی از تحقیقات گذشته مورد مطالعه بوده است. Frasson در سال ۱۹۹۶ مولفه های اصلی سیستم آموزشی هوشمند (مدل دانشجو، دانش دامنه و مدل آموزشی) را به شکل عامل های هوشمند طراحی کرده است [۶]. Lelouche در سال ۲۰۰۰ مجموعه ای از عامل های تعاملی را برای مدلسازی اصلی دانش آموزشی در سیستم آموزشی هوشمند بکار می گیرد [۷]. Legaspi و Sison در سال ۲۰۰۰ آموزش را در سیستم آموزشی هوشمند با استفاده از یادگیری تقویتی به عنوان رویه یادگیری مرکزی مدل کرده اند [۸]. Beck در سال ۱۹۹۸ یادگیری تقویتی را برای یادگیری اعمال تدریسی ممتاز وابسته با وضعیت های معین از دانش دانشجو استفاده کرده است [۹]. Baffes و Mooney در سال ۱۹۹۶، ASSERT را پیاده سازی کردند که یادگیری تقویتی را در مدلسازی دانشجو برای بدست آوردن خطاهای دانشجو جدید فقط با استفاده از دانش دامنه صحیح استفاده می کند [۱۰]. hashem در سال ۲۰۰۷ استفاده از اتوماتای یادگیر را در سیستم آموزشی هوشمند بکار گرفته است [۱۱].

مدلسازی دانشجو نقطه تمرکز اصلی و یکی از اجزاء سیستم آموزشی است. این مدل، نمایشی از رفتار و وضعیت دانشجو بوده و مبنایی برای نمایش وضعیت دانشجو است. هنگام استفاده دانشجو از سیستم، مدل دانشجو دنباله ای از اعمال انجام داده دانشجو را ثبت کرده و با استفاده از این سیستم الگوی یادگیری بکارگرفته شده توسط دانشجو را مدل کرده و پیشرفت یادگیری دانشجو را شناسایی می کند. هدف این مقاله بهبود مدلسازی دانشجو در سیستم یادگیری می باشد و نشان داده شده است که چگونه با استفاده از اتوماتای یادگیر و در چارچوب سیستم شبه آموزشی، می تواند این مدلسازی موفقیت آمیز باشد. سیستم آموزشی برای بالابردن منحنی یادگیری دانشجو نیاز دارد تا با هر دانشجو بر اساس مهارت و قابلیتش رفتار کند. به عبارتی دیگر این مقاله دیدگاه جدیدی را برای اینکه فرایند یادگیری دانشجو در سیستم شبه آموزشی مبتنی بر تئوری اتوماتای یادگیر تصادفی مدل شود، بیان می کند. اتوماتای یادگیر تصادفی سیستم را برای تشخیص مدل یادگیری دانشجو توانا می سازد و در نتیجه به سیستم شبه آموزشی جهت بهبود روش رفتار (مانند تدریس و ارزیابی) با دانشجو متناسب با مهارت او کمک می کند. برای بهبود تعیین مدل دانشجو با توجه به الگوریتم پیشنهادی به نام DPRI<sup>۲</sup> به جای الگوریتم DLRI<sup>۳</sup> در مدل Meta-LA [۱۱] استفاده شده است که با توجه به آزمایش های مختلف نشان داده شده است مدل بهبود یافته با دقت بیشتری مدل دانشجو را تشخیص می دهد. در صورتی که استاد بتواند مدل یادگیری دانشجو را تشخیص دهد، می تواند استراتژی تدریسش را با مدل یادگیری دانشجو منطبق کند. برای مثال اگر استاد بداند که با یک دانشجو ز رنگ سروکار دارد، آنگاه مسائل با پیچیدگی و سرعت بیشتر را برای دانشجو ارائه می دهد. در مدل پیشنهادی فرض بر این است که دانشجو می تواند به سه صورت دانشجوی کند، دانشجوی نرمال و دانشجوی سریع مدل شود. سیستم، مدل یادگیری دانشجو را با بررسی اعمال متوالی دانشجو و با بکارگیری Meta-LA تعیین می کند [۱۱ و ۱۲].

برای ارزیابی مدل مطرح شده از محیط های مختلفی شامل محک های ثابت تست استفاده شد و نتایج بدست آمده از شبیه سازی نشان می دهند که Meta-LA در تعیین مدل یادگیری بر اساس مشاهده اعمال دانشجو و محیط تدریس موفق بوده است. در ادامه مقاله در بخش ۲ به مفهوم اتوماتای یادگیری و انواع آن پرداخته شده است. در بخش ۳ سیستم شبه آموزشی توضیح داده شده است و در بخش ۴ مدل پیشنهادی بهبود یافته ارائه شده است. بخش ۵ به ارزیابی مدل پیشنهادی پرداخته و در بخش ۶ نتیجه گیری آورده شده است.

## ۲. اتوماتای یادگیر

اتوماتای یادگیر یکی از مدل های یادگیری تقویتی است که در آن اتوماتا یک عمل<sup>۴</sup> بهینه را از میان یک مجموعه از عمل های مجاز اتوماتا انتخاب می کند. این عمل روی یک محیط تصادفی اعمال می شود و محیط به این عمل اتوماتا بوسیله یک پاسخ تصادفی از مجموعه پاسخ های مجاز جواب می دهد. پاسخ محیط بصورت آماری به عمل اتوماتا وابسته است. اصطلاح محیط شامل تمام شرایط خارجی و تاثیرات آنها روی عملکرد اتوماتا است. ساختار یک اتوماتای یادگیر با محیط در شکل ۱ نشان داده شده است.



شکل ۱: اتصال اتوماتای یادگیر با محیط [۱۵]

محیط را می توان توسط سه تایی  $E \equiv \{\alpha, \beta, c\}$  نشان داد که در آن  $\alpha \equiv \{\alpha_1, \alpha_2, \dots, \alpha_r\}$  مجموعه ورودی ها،  $\beta \equiv \{\beta_1, \beta_2, \dots, \beta_m\}$  مجموعه خروجی ها و  $c \equiv \{c_1, c_2, \dots, c_r\}$  مجموعه احتمال های جریمه می باشد. هرگاه  $\beta$  مجموعه دو عضوی باشد، محیط از نوع P می باشد، در چنین محیطی  $\beta_1 = 1$  به عنوان جریمه و  $\beta_2 = 0$  به عنوان پاداش در نظر گرفته می شود. در محیط از نوع Q،  $\beta(n)$  می تواند به طور گسسته یک مقدار از مقادیر محدود در فاصله  $[0, 1]$  و در محیط از نوع S،  $\beta(n)$  متغیر تصادفی در فاصله  $[0, 1]$  است.  $C_t$  احتمال اینکه عمل  $\alpha_i$  نتیجه نامطلوب<sup>۵</sup> داشته باشد، می باشد. در محیط ایستای<sup>۶</sup> مقادیر  $C_t$  بدون تغییر می مانند، حال آن که در محیط غیر ایستای<sup>۷</sup> این مقادیر در طی زمان تغییر می کنند. اتوماتای یادگیر به دو گروه با ساختار ثابت و با ساختار متغیر تقسیم می گردد [۱۳ و ۱۴ و ۱۵]. اتوماتای یادگیر با ساختار ثابت با احتمالات گذر وضعیت ثابت مشخص<sup>۸</sup> می شود. نظریه زنجیره های مارکوف ابزار اصلی تجزیه و تحلیل این کلاس از اتوماتا می باشد و در اغلب موارد، رفتار مقتضی با انتخاب احتمالات گذر وضعیت اتوماتا در پاسخ به خروجی محیط به دست می آید.

اتوماتای یادگیر با ساختار متغیر<sup>۹</sup> توسط ۴ تایی  $\{\alpha, \beta, p, T\}$  نشان داده می شود که در آن  $\alpha \equiv \{\alpha_1, \alpha_2, \dots, \alpha_r\}$  مجموعه عمل های اتوماتا،  $\beta \equiv \{\beta_1, \beta_2, \dots, \beta_m\}$  مجموعه ورودی های اتوماتا و  $p = \{p_1, p_2, \dots, p_n\}$  بردار احتمال انتخاب هر یک از اعمال و  $p(n+1) = T[\alpha(n), \beta(n), p(n)]$  الگوریتم یادگیری می باشد. در این نوع از اتوماتاها، اگر عمل  $\alpha_i$  در مرحله n ام انتخاب شود و این عمل، پاسخ مطلوب از محیط دریافت نماید، احتمال  $p_i(n)$  افزایش یافته و سایر احتمالات کاهش می یابند. برای پاسخ نامطلوب احتمال  $p_i(n)$  کاهش یافته و سایر احتمالات افزایش می یابند. در هر حال، تغییرات به گونه ای صورت می گیرد تا حاصل جمع  $p_i(n)$  ها همواره ثابت و مساوی یک باقی بماند. الگوریتم زیر بر اساس (۲) و (۱) نمونه ای از الگوریتم های یادگیری خطی در اتوماتای با ساختار ثابت است [۱۵].

الف- پاسخ مطلوب برای عمل  $\alpha_i$ :

$$\begin{aligned} p_i(n+1) &= p_i(n) + a[1 - p_i(n)] \\ p_j(n+1) &= (1-a)p_j(n) \quad \forall j \neq i \end{aligned} \quad (1)$$

ب- پاسخ نامطلوب برای عمل  $\alpha_i$ :

$$\begin{aligned} p_i(n+1) &= (1-b)p_i(n) \\ p_j(n+1) &= (b/r - 1) + (1-b)p_j(n) \quad \forall j \neq i \end{aligned} \quad (2)$$

در روابط فوق،  $a$  پارامتر پاداش و  $b$  پارامتر جریمه می باشد. با توجه به مقادیر  $a$  و  $b$  سه حالت زیر را می توان در نظر گرفت. زمانی که  $a$  و  $b$  با هم برابر باشند، الگوریتم را LRP<sup>۱۰</sup> می نامیم. زمانی که  $b$  از  $a$  خیلی کوچکتر باشد، الگوریتم را LREp<sup>۱۱</sup> می نامیم و زمانی که  $b$  مساوی صفر باشد، الگوریتم را LRI<sup>۱۲</sup> می نامیم.

اتوماتاهای یادگیر دارای کاربردهای فراوانی می باشند. بعضی از این کاربردها عبارتند از: مسیریابی در شبکه های ارتباطی، فشردن سازی تصاویر، شناسایی الگو، زمانبندی فرایندها در شبکه های کامپیوتری، تئوری صف، کنترل دسترسی در شبکه های انتقال ناهمزمان، کمک به آموزش شبکه های عصبی، دسته بندی و افراز اشیاء و پیدا کردن ساختار بهینه برای شبکه های عصبی [۱۶ و ۱۷ و ۱۸].

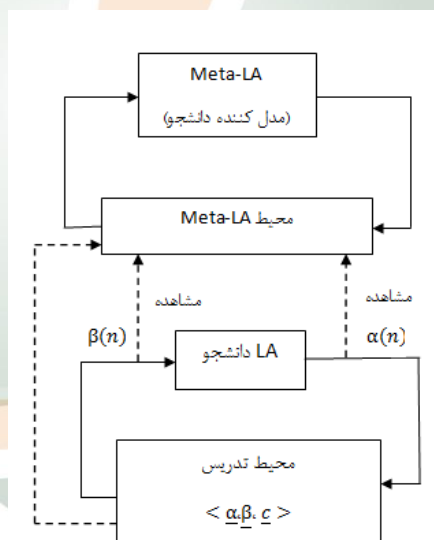
### ۳. سیستم شبه آموزشی<sup>۱۳</sup>

سیستم های شبه آموزشی نمایشی از وضعیت دانشجو می باشند. این سیستم شبه آموزشی، توانایی کارکردن و تست بدون نیاز به اشخاص واقعی را دارا می باشد. یکی از مشکلات هر سیستم real-life با دانشجویان real-life این است که سیستم باید برای رسیدن به یک استراتژی تدریس بهینه از قبل توسط دانشجویان بسیاری، تست شود. استفاده از سیستم شبه ساز دانشجو، بطور امیدوارکننده ای، این مشکل را برطرف می کند و سیستم شبه آموزشی را برای رسیدن به استراتژی بهینه بدون نیاز به اشخاص برای تعامل با یک سیستم آموزشی بهینه، توانا می سازد [۱۲].

در اینجا از شبه ساز دانشجویی که اتوماتای یادگیر تصادفی را بعنوان یک مکانیزم یادگیری برای تقلید اینکه دانشجو چگونه رفتار می کند و در یک قسمت آموزشی چگونه یاد می گیرد، استفاده می شود. فرض بر این است که به کارگیری اتوماتای یادگیری اجازه خواهد داد مدل شبه ساز خصوصیتی از دانشجوی واقعی را تشخیص داده و به همان اندازه تعاملات دانشجو با سیستم شبه آموزشی را افزایش دهد و همچنین یادگیری از اتوماتای یادگیری افزایش می یابد، که بر این دلالت دارد که مدل شبه ساز، یادگیری بهتری از دانشجو ارائه می کند. این انتظار وجود دارد که بعد از تکمیل سیستم، مدل شبه ساز دانشجو می تواند توسط دانشجویان واقعی جایگزین شود [۱۲].

#### ۳-۱. ساختار شبکه اتوماتای یادگیر Meta LA/LA در سیستم شبه آموزشی

سیستم شبه آموزشی از چند اتوماتای یادگیر تشکیل شده است که بطور غیر مستقیم به همدیگر وصل شده اند (یک سیستم اتوماتای بهم متصل است). نحوه اتصال اتوماتاهای یادگیر به دو صورت همزمان<sup>۱۴</sup> یا ترتیبی<sup>۱۵</sup> می باشد که این سیستم از مدل همزمان استفاده می کند. چگونگی اتصالات اتوماتای یادگیر در این سیستم در شکل ۲ نشان داده شده است [۱۱ و ۱۲ و ۱۹ و ۲۰].



شکل ۲: مدل سازی با استفاده از شبکه ای از اتوماتاهای یادگیر [۱۱]

در اتوماتای متصل سنتی پاسخ محیط یک اتوماتا، ورودی اتوماتای دیگر را بیان می کند [۱۵]. در حالی که در این سیستم، تا زمانیکه اتوماتای یادگیر دانشجویان روی Meta-LA تاثیر گذار باشد، مابین آنها ارتباط مستقیمی وجود ندارد. محیط Meta-LA کارایی اتوماتای یادگیر دانشجو را در یک دوره زمانی کنترل کرده و اینکه به عمل انتخابی، Meta-LA پاداش یا جریمه دهد، به محیط آن بستگی دارد. این مدل یک ساختار جدید از ارتباطات داخلی را نشان می دهد که ترکیبی از دو سطح است. یک اتوماتای سطح بالا مثل Meta-LA و یک اتوماتای سطح پایین که اتوماتای یادگیر دانشجو است. همگرایی اتوماتای سطح بالا به رفتار اتوماتای سطح پایین بستگی دارد. یک سناریو نمونه می گوید که اتوماتای سطح بالا می تواند قبل از همگرایی اتوماتای سطح پایین، همگرا شود در حالیکه اگر اتوماتای سطح پایین قبل از اتوماتای سطح بالا همگرا شود، این به معنای همگرایی نادرست اتوماتای سطح بالا است.

از مشاهده تعامل مازول ها در شکل نتایجی که در ادامه بیان شده است استنتاج می شود: محیط Meta-LA ترکیبی از مشاهدات است، که باید بطور ضمنی بعد از انجام پردازش پیام های مشاهده شده، دریافت شود، Meta-LA به محیط سطح پایین که شامل مجموعه ای از احتمالات جریمه می باشد و برای اتوماتای یادگیر سطح پایین ناشناخته است، دسترسی دارد، Meta-LA به اعمال انتخاب شده توسط اتوماتای یادگیر سطح پایین و نیز پاسخ های ایجاد شده توسط محیط سطح پایین دسترسی دارد. در مجموع، همانطور که در شکل مشخص می باشد، سه نتیجه انتهایی، مجموعاً محیط Meta-LA را تشکیل می دهند [۱۱ و ۱۲ و ۱۹ و ۲۰].



#### ۴. مدل پیشنهادی و بهبود مدل سازی دانشجو

سیستم شبه آموزشی یک مدل را برای دانشجو توسط ارزیابی و تعیین راهی که او یاد می گیرد، استنتاج می کند. برای رسیدن به این، فرض شده که سیستم یک مجموعه متناهی از مدل های یادگیری ممکن را برای هر دانشجو دارد. مدل دانشجو تقریباً، نمایش کیفی از دانش دانشجو در مورد یک دامنه خاص، یا موضوع خاص یا مهارت در دامنه است، که می تواند کاملاً یا بطور جزئی برای وجه های خاصی از رفتار دانشجو محاسبه شود. دامنه دانش با سوالات مدل سقراطی<sup>۱۶</sup> چند گزینه ای نمایش داده می شود و بنابراین در هر سوال، به هر گزینه احتمال پاداشی وابسته شده است و گزینه با بیشترین احتمال پاداش، پاسخ سوال در نظر گرفته شده است [۲۱].

فرض شده است هر دانشجو می تواند با یکی از سه نوع، یک مدل برگرفته از اتوماتای یادگیر با ساختار ثابت که آن را  $FSSA^{17}$  می نامیم برای یک دانشجوی کند، مدل مبتنی بر اتوماتای یادگیر متغیر  $VSSA^{18}$  برای دانشجوی نرمال و مدل مبتنی بر اتوماتای Pursuit برای دانشجوی سریع مدل شود. این مجموعه متناهی از مدل های یادگیری راه های مختلف یادگیری دانشجو را ارائه می دهد. اگر سیستم شبه آموزشی بتواند چگونگی درک دانش دانشجو را بفهمد، آنگاه می تواند استراتژی تدریس بهینه را برای هر دانشجو پیدا کند. برای مثال، اگر سیستم شبه آموزشی تشخیص دهد که دانشجوی در حال یادگیری، مشابه مدل Pursuit است، آنگاه می تواند نتیجه بگیرد که دانشجو زنگ است و اینکه آن یک یادگیرنده سریع می باشد. در این حالت سیستم می تواند استراتژی تدریسش را برای آمادگی یک دانشجوی سریع انطباق دهد که آن را برای بهبود استراتژی تدریسش توانا می سازد.

مدل دانشجو از سه نوع خاص اتوماتای یادگیری استفاده خواهد کرد، هریک از اینها به صورت یک عمل برای اتوماتای سطح بالا مورد استفاده قرار می گیرند. مدل اتوماتای یادگیر یک مدل با شمای پیوسته است که، در این حالت احتمالات عمل می توانند هر مقدار واقعی در بازه  $[0,1]$  باشند. Thathachar و Oommen مفهوم اتوماتای یادگیر گسسته را توسط محدود کردن احتمال انتخاب از یک مجموعه متناهی از مقادیر در بازه بسته  $[0,1]$  مطرح کردند. بنابراین احتمال عمل در گام ها به صورت گسسته برورسانی می شود به جای آنکه برورسانی در حالت پیوسته باشد [۲۲]. مهمترین مزیت استفاده از الگوریتم های گسسته افزایش سرعت همگرایی است [۲۳]. مزیت دیگر استفاده از این الگوریتم ها حذف شدت اعمال نفوذ بر روی مولد اعداد تصادفی است و مزیت سوم آن، استفاده کمتر از حافظه و اجرای سریعتر از اتوماتای پیوسته است. Oommen اثبات کرد که اتوماتاهای گسسته در همه محیط ها نزدیک به بهینه بوده و از همتاهای پیوسته شان همگرایی سریعتری دارند [۲۴]. در روش پیشنهادی از الگوریتم های تعقیبی استفاده شده است که درواقع، کلاسی از الگوریتم های تخمین زننده است و عمل بهینه را پیگیری می کند و شامل سه مرحله است، مرحله اول عملی بر مبنای توزیع احتمال  $P(t)$  انتخاب شده و در مرحله دوم اگر اتوماتا پاداش بگیرد، مولفه ای از  $p(t)$  که تخمین پاداش جاری اش ماکزیمم است افزایش می یابد و احتمالات عمل بقیه کاهش می یابد و نهایتاً در مرحله سوم تخمین های جاری احتمال برورسانی می شود. این الگوریتم را  $DP_{RI}$  می نامیم. به نظر می رسد الگوریتم  $DP_{RI}$  باید همگرایی سریعتری نسبت به  $DL_{RI}$  داشته باشد و در نهایت مدل سازی دانشجو بهبود داده می شود.

در مثال ارائه شده Meta-LA برای یادگرفتن بهترین مدل ممکن برای شبیه سازی دانشجو از چهارتایی  $\{\alpha, \beta, P, T\}$  استفاده می کند که در آن مجموعه اعمال به صورت  $\alpha = \{\alpha_1, \alpha_2, \alpha_3\}$  است.  $\alpha_1$  نشان می دهد عمل با  $FSSA$ ، یادگیرنده کند مطابق است،  $\alpha_2$  نشان می دهد عمل با  $VSSA$ ، یادگیرنده نرمال مطابق است و  $\alpha_3$  نشان می دهد عمل با Pursuit، یادگیرنده سریع مطابق است.  $\beta = \{0,1\}$ ،  $\beta = 0$  اشاره می کند که پاداش برای عمل حاضر انتخاب شده است و  $\beta = 1$  اشاره می کند که جریمه برای عمل حاضر انتخاب شده است.  $P = [p_1, p_2, p_3]^T$ ،  $p_i(n)$  احتمالات جاری شبیه ساز دانشجو که با  $\alpha_i$  نشان داده شده است و  $T$  قانون برورسانی احتمالات است که بصورت مقابل نگاشت می شود:  $T: (P, \alpha, \beta) \rightarrow P$ . هریک از اینها در ادامه بصورت کامل توضیح داده شده اند:

$\beta$  ورودی است که Meta-LA دریافت می کند. Meta-LA باید یک دنباله ای از تصمیمات اخذ شده توسط شبیه ساز دانشجو را مشاهده کند و بر مبنای این دنباله باید نتیجه گیری کند که آیا مدل جاری اش درست است یا نه؟ برای رسیدن به این پیشنهاد شده است که، برای تعداد ثابت پرس و جوها، Meta-LA فرض کند مدل شبیه ساز دانشجو  $\alpha_i$  است. فرض می کنیم که  $\theta(t)$  نرخ یادگیری اتوماتای یادگیر سطح پایین در زمان  $t$  باشد. توسط بازبینی راهی که شبیه ساز دانشجو در مدت این پرس و جوها یاد می گیرد، استنتاج می شود که آیا مدل جاری باید پاداش بگیرد یا جریمه شود؟ ما توسط دو آستانه به هدف مان می رسیم:  $\{ \theta_j | 1 \leq j \leq 2 \}$  اگر نرخ یادگیری مشاهده شده جاری از شبیه ساز دانشجو کمتر از  $\theta_1$  باشد، ادعا می کنیم که شبیه ساز دانشجو یادگیرنده سریع است. حال اگر Meta-LA،  $\alpha_1$  را انتخاب کند، سیستم پاداش را انتخاب می کند در غیر اینصورت جریمه انتخاب می شود.

اگر نرخ یادگیری مشاهده شده جاری از شبیه ساز دانشجو بزرگتر مساوی  $\theta_1$  باشد و یا کمتر از  $\theta_2$ ، ادعا می کنیم که شبیه ساز دانشجو یادگیرنده نرمال است. حال اگر Meta-LA،  $\alpha_2$  را انتخاب کند، سیستم پاداش را انتخاب می کند در غیر اینصورت جریمه انتخاب می شود. اگر نرخ یادگیری مشاهده شده جاری از شبیه ساز دانشجو مساوی یا بزرگتر از  $\theta_2$  باشد، ادعا می کنیم که شبیه ساز دانشجو

خصوصیات یادگیرنده کند را دارد. حال اگر Meta-LA،  $\alpha_3$  را انتخاب کند، سیستم پاداش را انتخاب می کند در غیر اینصورت جریمه انتخاب می شود.  $p$  بردار احتمال عمل است که شامل احتمالاتی است که Meta-LA به هر یک از اعمالش نسبت می دهد. در لحظه  $t$ ، Meta-LA بطور تصادفی یک عمل  $\alpha_i(t) = \alpha_i$  را انتخاب می کند. احتمالی که آتوماتا عمل  $\alpha_i$  را در زمان  $t$  انتخاب کند، احتمال عمل است و در هر حال، تغییرات به گونه ای صورت می گیرد تا حاصل جمع  $p_i(n)$  ها همواره ثابت و مساوی یک باقی بماند که در (۳) نشان داده شده است.

جدول ۱ آستانه مورد استفاده توسط Meta-LA برای مدل یادگیری [۱۱]

حد آستانه	انواع دانشجو
$\theta$	یادگیرنده سریع
$\theta$	یادگیرنده نرمال
$\theta$	یادگیرنده کند

$$p_i(t) = pr[\alpha(t) = \alpha_i], \quad (۳)$$

$$where, \sum_{i=1}^r p_i(t) = 1, \forall t, i = 1, 2, 3$$

در آغاز Meta-LA، یک احتمال مساوی برای هر یک از سه مدل یادگیری در نظر می گیرد که در (۴) بیان شده است:

$$p(0) = [0.333, 0.333, 0.333]^T \quad (۴)$$

$T$ ، الگوریتم بروز رسانی یا شمای کمکی است که برای بروز رسانی بردار احتمال عمل Meta-LA بکارگرفته می شود و بصورت (۵) نمایش داده می شود:

$$p(t+1) = T[p(t), \alpha(t), \beta(t)] \quad (۵)$$

این الگوریتم بروز رسانی می تواند هر یک از شمای یادگیری LA باشد. در روش پیشنهادی از شمای DPRI استفاده شده است که در این شما هنگامی که یک عمل پاداش می گیرد، همه اعمالی که منطبق بر بیشترین تخمین نیستند، کاهش می یابند و عمل با بیشترین تخمین افزایش می یابد. مجموع بردار احتمال همیشه برابر با یک است. هنگامی که عمل انتخابی جریمه می شود، بردار احتمال عمل بروز رسانی نمی شود [۲۵]. شمای DPRI در (۶) بیان شده است.

$$(۶) \quad \text{Else} \quad p_j(t+1) = p_j(t) \text{ for all } 1 \leq j \leq r$$

Meta-LA به عنوان یک آتوماتای یادگیر سطح بالا پیاده سازی شده است و همانطور که در قسمت ۳-۱ بیان شده است به محیط یادگیری، اعمال و جریمه های آتوماتای یادگیر سطح پایین وابسته به شبیه ساز دانشجو دسترسی دارد بطوریکه بعدا مواد تدریسی را یاد می گیرد. Meta-LA به عنوان یک آتوماتای یادگیر گسسته با شمای DPRI پیاده سازی شده است. شبیه ساز دانشجو برای تقلید سه نوع از دانشجویان بصورت زیر شبیه سازی شده است:

- یادگیرنده کند: برای این نوع از دانشجویان، شبیه ساز دانشجو از FSSA برای تقلید رفتار دانشجو استفاده می کند. در اصل از آتوماتای Tsetlin با دو وضعیت برای هر عمل استفاده می کند.
- یادگیرنده نرمال: برای این گروه از دانشجویان، شبیه ساز دانشجو برای شبیه سازی رفتار دانشجو از VSSA استفاده می کند که در این مقاله با شمای  $L_{RI}$  و با  $\mu = 0.05$  پیاده سازی شده است.
- یادگیرنده سریع: برای شبیه سازی دانشجو از این نوع، شبیه ساز دانشجو از آتوماتای یادگیر Pursuit استفاده می کند که در این مقاله با شمای  $PL_{RI}$  و با  $\mu = 0.05$  پیاده سازی شده است.

□، نرخ یادگیری دانشجو، محیط Meta-LA از فرمول (۷) استفاده می کند:

$$\frac{1}{n} \sum \mathbf{n}_i \mathbf{c}_i \quad (\gamma)$$

که  $n_i$  تعداد دفعاتی است که عمل  $\alpha_i$  در طول انجام یک تعداد ثابت از پرس و جوها انتخاب می شود و  $C_i$  احتمال جریمه وابسته با عمل  $\alpha_i$  اتوماتای یادگیر سطح پایین است. نرخ یادگیری دانشجو تابعی از زمان است یعنی با گذشت زمان این مقدار برای هر سه نوع دانشجو کاهش می یابد زیرا دانش دانشجو با گذشت زمان افزایش می یابد.

## ۵. ارزیابی مدل

در این قسمت نتایج بدست آمده از شبیه سازی مدل پیشنهادی را نشان می دهیم. برای بدست آوردن این نتایج شبیه سازی های مختلفی را اجرا نمودیم تا مشخص نماییم مدل کننده دانشجو تا چه اندازه می تواند مدل یادگیری دانشجو را درست تشخیص دهد و نتایج بدست آمده با نتایج مدل قبلی [۱۱] مقایسه شده اند. شبیه سازی ها در انواع مختلفی از محیط ها تست شده اند که این محیط ها مسئله تدریس را نشان می دهند که شامل مواد درسی نامعلوم از مدل دامنه وابسته با سیستم شبه آموزشی می باشد. در همه آزمایش های اجرا شده، اگر احتمال انتخاب یک عمل بزرگتر یا مساوی به آستانه  $T (0 < T \leq 1)$  شود، الگوریتم همگرا در نظر گرفته می شود. اگر آتوماتا به بهترین عمل (عملی با بیشترین احتمال پاداش) همگرا شود آن را همگرایی صحیح بیان می کنیم.

شبیه سازی ها برای محیط های محک موجود تست شدند و حد آستانه، ۰,۹۹، تنظیم شد و ۷۵ بار آزمایشات اجرا شد و نتیجه، میانگین ۷۵ اجرا می باشد. در شبیه سازی ها فرض شده است که هنگامی که دانشجو استفاده از سیستم را آغاز می کند، نتایج در طول یک فاز یادگیری کوتاه نادرست است و بنابراین ۲۰ تکرار اول (تمام تکرارها، تصمیم گیری ها و پاسخ های اتوماتای سطح پایین) توسط Meta-LA صرف نظر می شود. نتایج این شبیه سازی ها در ادامه توضیح داده می شود.

### الف) آزمایش در محیط با ۴ عمل

محیط با ۴ عمل یک سوال ۴ گزینه ای را نشان می دهد. دانشجو باید محتویات این سوال را یاد بگیرد. بر اساس کارایی دانشجو -Meta-LA یک مدل یادگیری برای دانشجو بدست می آورد. دو تنظیم مختلف برای محیط ها به همراه احتمال های پاداش در جداول مشخص شده اند. نتایج بدست آمده در محیط EA که بهترین عمل  $\alpha_1$  می باشد در جدول ۲ نشان داده شده است. در زمانیکه اتوماتای سطح پایین تر از الگوریتم pursuit استفاده می کند، در الگوریتم پیشنهادی ما Meta-LA قادر به تشخیص مدل یادگیری درست دانشجو در ۹۱٪ زمان می باشد، در حالیکه این مقدار در الگوریتم پیشنهادشده در مرجع [۱۱]، ۷۶٪ می باشد. همچنین در زمان استفاده از اتوماتای L<sub>RI</sub>، درصد همگرایی صحیح در الگوریتم ما ۹۴٪ و در الگوریتم پیشنهادشده در مرجع [۱۱]، ۷۴٪ و همینطور در زمان استفاده اتوماتای سطح پایین تر از Tsetlin، این مقادیر به ترتیب ۱۰۰٪ و ۵۳٪ می باشند.

نتایج بدست آمده در محیط EB که بهترین عمل  $\alpha_3$  می باشد و پیچیدگی بیشتری نسبت به محیط EA دارد در جدول ۳ نشان داده شده است. همانطور که در جدول مشخص شده است، الگوریتم پیشنهادی در تعیین دانشجو کند و نرمال در ۱۰۰٪ مواقع موفق بوده و در تعیین دانشجوی سریع در ۹۷٪.

جدول ۲ مقایسه همگرایی Meta\_LA در مدل پیشنهادی و مدل پیشنهادشده در مرجع [۱۱] در محیط E<sub>(A,F)</sub>

$E_{(A,4)}$ محیط			$PL_{RI}$		$L_{RI}$		Tsetlin	
			تعداد تکرار	همگرایی صحیح	تعداد تکرار	همگرایی صحیح	تعداد تکرار	همگرایی صحیح
مدل پیشنهادشده در مرجع [۱۱]	۰٫۴۶ – ۰٫۳۵	۰٫۵۷ – ۰٫۴۰	۵۰۶	%۷۶	۷۱۷	%۷۴	۶۸۲	%۵۳
مدل پیشنهادی			۵۹۹	%۹۱	۵۴۲	%۹۴	۱۴۲۳	%۱۰۰
احتمال پاداش	$E_A = \{0,7 \quad 0,5 \quad 0,3 \quad 0,2\}$							

جدول ۳ مقایسه همگرایی Meta\_LA در مدل پیشنهادی و مدل پیشنهادشده در مرجع [۱۱] در محیط  $E_{(B,F)}$

محیط $E_{(B,F)}$			$PL_{RI}$		$L_{RI}$		Tsetlin	
			تعداد تکرار	همگرایی صحیح	تعداد تکرار	همگرایی صحیح	تعداد تکرار	همگرایی صحیح
مدل پیشنهادشده در مرجع [۱۱]	۰,۲۸ - ۰,۱۷	۰,۳۷ - ۰,۲۴	۷۹۸	%۶۹	۹۲۴	%۷۶	۴۲۳	%۸۴
مدل پیشنهادی			۸۲۹	%۹۷	۹۷۹	%۱۰۰	۸۴۳	%۱۰۰
احتمال پاداش	$E_B=\{ \quad ۰,۱ \quad \quad ۰,۴۵ \quad \quad ۰,۸۴ \quad \quad ۰,۷۶ \}$							

### ب) آزمایش در محیط با ۱۰ عمل

محیط با ۱۰ عمل، یک سوال با ۱۰ گزینه را نمایش می دهد. احتمالات پاداش محیط در جدول ۴ مشخص شده است. نتایج بدست آمده در محیط EA که بهترین عمل  $\alpha_1$  می باشد در جدول ۴ نشان داده شده است. در زمانیکه اتوماتای سطح پایین تر از PL<sub>RI</sub> استفاده می کند، در الگوریتم پیشنهادی ما Meta-LA قادر به تشخیص مدل یادگیری درست دانشجو در %۹۷ زمان می باشد، در حالیکه این مقدار در الگوریتم پیشنهادشده در مرجع [۱۱]، %۹۲ می باشد. همچنین در زمان استفاده از اتوماتای L<sub>RI</sub>، درصد همگرایی صحیح در الگوریتم ما %۱۰۰ و در الگوریتم پیشنهادشده در مرجع [۱۱]، %۱۰۰ بوده است که در این حالت الگوریتم ما با تعداد تکرار کمتری به این نتیجه رسیده است و همینطور در زمان استفاده از اتوماتای سطح پایین تر از Tsetlin، این مقادیر به ترتیب %۱۰۰ و %۹۳ می باشند.

نتایج بدست آمده در محیط EB که بهترین عمل  $\alpha_3$  می باشد و پیچیدگی بیشتری نسبت به محیط EA دارد در جدول ۵ نشان داده شده است. همانطور که در جدول مشخص شده است، الگوریتم پیشنهادی در تعیین دانشجو کند و نرمال در %۱۰۰ مواقع موفق بوده و در تعیین دانشجوی سریع در %۹۴.

جدول ۴ مقایسه همگرایی Meta\_LA در مدل پیشنهادی و مدل پیشنهادشده در مرجع [۱۱] در محیط  $E_{(A,10)}$

محیط $E(A,10)$			PL <sub>RI</sub>		L <sub>RI</sub>		Tsetlin	
			تعداد	همگرایی	تعداد	همگرایی	تعداد	همگرایی
			تکرار	صحیح	تکرار	صحیح	تکرار	صحیح
مدل پیشنهادشده در مرجع [۱۱]	۰,۴۶ – ۰,۳۵	۰,۵۷ – ۰,۴۰	۵۴۴	%۹۲	۴۹۷	%۱۰۰	۶۱۰	%۹۳
مدل پیشنهادی			۵۳۰	%۹۷	۴۵۸	%۱۰۰	۱۹۰۶	%۱۰۰
احتمال پاداش	E <sub>A</sub> = { ۰,۷   ۰,۵   ۰,۳   ۰,۲   ۰,۴   ۰,۵   ۰,۴   ۰,۳   ۰,۵   ۰,۲ }							

جدول ۵ مقایسه همگرایی Meta\_LA در مدل پیشنهادی و مدل پیشنهادشده در مرجع [۱۱] در محیط  $E_{(B,10)}$

محیط E(B,۱۰)			PL <sub>RI</sub>		L <sub>RI</sub>		Tsetlin	
			تعداد تکرار	همگرایی صحیح	تعداد تکرار	همگرایی صحیح	تعداد تکرار	همگرایی صحیح
مدل پیشنهادشده در مرجع [۱۱]	۰,۲۸ - ۰,۱۷	۰,۳۷ - ۰,۲۴	۸۷۰	%۶۸	۱۰۰۷	%۷۷	۴۳۴	%۱۰۰
مدل پیشنهادی			۱۱۳۹	%۹۴	۸۶۸	%۱۰۰	۴۳۴	%۱۰۰
احتمال پاداش	E <sub>B</sub> ={۰,۱    ۰,۴۵    ۰,۸۴    ۰,۷۶    ۰,۲    ۰,۴    ۰,۶    ۰,۷    ۰,۵    ۰,۳ }							

### ۶. جمع بندی

این مقاله استراتژی جدیدی برای سیستم شبه آموزشی بیان می کند که به مدلسازی دانشجو اشاره دارد. مدل کننده دانشجو خودش از اتوماتای یادگیر به عنوان یک مکانیزم درونی برای تعیین مدل یادگیری دانشجو استفاده می کند بطوریکه می تواند در سیستم شبه آموزشی برای تعیین تمرین یادگیری برای هر دانشجو مورد استفاده قرار گیرد. برای رسیدن به این، مدل کننده دانشجو از اتوماتای سطح بالاتر به نام Meta-LA استفاده می کند که اعمال اتوماتاهای شبیه ساز دانشجو و محیط تدریس را مشاهده می کند و تلاش می کند تا مدل یادگیری دانشجو را تعیین کند. Meta-LA مشخص می کند که دانشجو در سوال یک یادگیرنده کند، سریع و یا نرمال است.

این مقاله با استفاده از بکارگیری الگوریتم pursuit دقت همگرایی و در نتیجه تعیین مدل یادگیری دانشجو را بهبود می دهد و به سیستم شبه آموزشی برای تعیین بهتر مدل دانشجو کمک می کند. نتایج شبیه سازی ها نشان داد که این روش مکانیزم معتبر و امکان پذیر برای



اجرای فرایند یادگیری دانشجو است و همچنین نشان داد که مدل دانشجو در تعیین مدل یادگیری دانشجو موفق است و استفاده از الگوریتم pursuit در درصد بالایی از حالت ها، بعضی اوقات با دقت ۱۰۰٪ مدل یادگیری را مشخص کرده است.

## ۷. مراجع

- [۱] V.J.Shute; J. Psotka; "Intelligent Tutoring Systems: Past, Present, and Future", Handbook of Research on Educational Communications and Technology, Scholastic Publications, ۱۹۹۵.
- [۲] E. Wenger; "Artificial Intelligence and Tutoring Systems: Computational and Cognitive Approaches to the Communication of Knowledge", Los Altos, CA: Morgan Kaufmann Publishers, Inc, ۱۹۸۷.
- [۳] E. Fischetti ; A. Gisolfi; "From computer-aided instruction to intelligent tutoring systems", Educ. Technol, vol, ۳۰, no. ۸, pp. ۷-۱۷, ۱۹۹۰.
- [۴] R. Winkels ; J. Breuker; "What's in an ITS? a functional decomposition", in New Directions for Intelligent Tutoring systems, E. Costa, Ed.Berlin, Germany: Springer-Verlag, ۱۹۹۰.
- [۵] J. Self; "The defining characteristics of intelligent tutoring systems research: ITSs care, precisely", Int. J. Artif. Intell. Educ, vol. ۱۰, pp. ۳۵۰-۳۶۴, ۱۹۹۹.
- [۶] C. Frasson; T. Mengelle; E. Aimeur; G. Gouarderes; "An actor-based architecture for intelligent tutoring system", In intelligent Tutoring system: ۳<sup>rd</sup> international Conference, ITS '۹۶: Lecture Notes in computer Science, pages ۵۷-۶۵. Springer-verlag, Berlin, ۱۹۹۶.
- [۷] R.Lelouche; "A collection of pedagogical agents for intelligent educational systems", In intelligent tutoring systems: ۶<sup>th</sup> international Conference, ITS ۲۰۰۰: Lecture Notes in computer Science, pages ۱۴۳-۱۵۲. Springer-verlag, Berlin, ۲۰۰۰.
- [۸] R. S. Legaspi ; R. C. Sison; "Modeling the tutor using reinforcement learning", in Proc. PCSC, pp. ۱۹۴-۱۹۶, ۲۰۰۰.
- [۹] J. Beck; "Learning to teach with a reinforcement learning agent", in Proc. ۱۵<sup>th</sup> Nat. Conf. AI/IAAI, Madison, WI, p. ۱۱۸۵, ۱۹۹۸.
- [۱۰] P. Baffes ; R. Mooney; "Refinement-based student modeling and automated bug library construction", J. Artif. Intell. Educ, vol. ۷, no. ۱, pp. ۷۵-۱۱۶, ۱۹۹۶.
- [۱۱] M.K., Hashem; B.J Oommen; "Modeling a Student's Behavior in a Tutorial-Like System Using Learning Automata ", IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics., ۱۰۸۳-۴۴۱۹/\$۲۶,۰۰ © ۲۰۰۹ IEEE.
- [۱۲] M. K. Hashem; "Learning automata based intelligent tutorial-like systems", Ph.D. dissertation, School Comput. Sci., Carleton Univ., Ottawa, ON, Canada, ۲۰۰۷.
- [۱۳] M. A. L. Thathachar ; P. S. Sastry; "Varieties of learning automata: An overview", IEEE Trans. Syst., Man, Cybern. B, Cybern., vol. ۳۲, no. ۶, pp. ۷۱۱-۷۲۲, Dec. ۲۰۰۲.
- [۱۴] S. Lakshmivarahan; "Learning Algorithms Theory and Applications" New York: Springer-Verlag, ۱۹۸۱.
- [۱۵] K. S. Narendra ; M. A. L. Thathachar; Learning Automata: An Introduction. Englewood Cliffs, NJ: Prentice-Hall, ۱۹۸۹.
- [۱۶] M. R. Meybodi; S. Lakshmivarhan; "A Learning Approach to Priority Assignment in a Two Class M/M/۱ Queuing System with Unknown Parameters", Proceedings of Third Yale Workshop on Applications of Adaptive System Theory, Yale University, ۱۰۶-۱۰۹, ۱۹۸۳.
- [۱۷] M. R. Meybodi ; H. Beigy; "New Class of Learning Automata Based Scheme for Adaptation of Backpropagation Algorithm Parameters", Proceedings of EUFIT-۹۸, Aachen, Germany, ۳۳۹-۳۴۴, ۷-۱۰ Sep. ۱۹۹۸.
- [۱۸] D. S. Johnson; L. A. McGeoch; "Experimental Analysis of Heuristics for the STSP", in the Traveling Salesman Problem and its Variations, G. Gutin and A. Punnen, Editors, Kluwer Academic Publishers, Boston, ۳۶۹-۴۴۳, ۲۰۰۲.
- [۱۹] M.K., Hashem; B.J. Oommen; "On using learning automata to model a student's behavior in a tutorial-like system", In: Okuno, H.G, Ali, M. (eds.) IEA/AIE ۲۰۰۷. LNCS (LNAI), vol. ۴۵۷۰, pp. ۸۱۳-۸۲۲, Springer, Heidelberg (۲۰۰۷).
- [۲۰] M.K., Hashem; B.J. Oommen; " Learning Automata Based Intelligent Tutorial-like System ", J.D. Velásquez et al. (Eds.): KES ۲۰۰۹, Part I, LNAI ۵۷۱۱, pp. ۳۶۱-۳۷۴, ۲۰۰۹. © Springer-Verlag Berlin Heidelberg ۲۰۰۹.
- [۲۱] M.K., Hashem; B.J. Oommen; "Using learning automata to model a domain in a tutorial-like system ", In: Proceedings of ICMLC ۲۰۰۷, the ۲۰۰۷ International Conference of Machine Learning and Cybernetics, Hong Kong, pp. ۱۱۲-۱۱۸, August ۲۰۰۷.
- [۲۲] M. A. L. Thathachar ; B. J. Oommen; "Discretized reward-inaction learning automata", J. Cybern. Inf. Sci., vol. ۲, no. ۱, pp. ۲۴-۲۹, Spring ۱۹۷۹.
- [۲۳] B. J. Oommen ; J. K. Lanctôt; "Discretized pursuit learning automata", IEEE Trans. Syst., Man, Cybern., vol. ۲۰, no. ۴, pp. ۹۳۱-۹۳۸, Jul./Aug. ۱۹۹۰.
- [۲۴] M. A. L. Thathachar ; B. J. Oommen; "Discretized reward-inaction learning automata", J. Cybern. Inf. Sci., vol. ۲, no. ۱, pp. ۲۴-۲۹, Spring ۱۹۷۹.
- [۲۵] M. Agache; B.J .Oommen; "Generalized pursuit learning schemes: New families of continuous and discretized learning automata ", IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics ۳۲(۶), ۷۳۸-۷۴۹ (۲۰۰۲).

<sup>۱</sup> Pedagogical

<sup>۲</sup> Discrete Pursuit Reward Inaction

<sup>۳</sup> Discrete Linear Reward Inaction

<sup>۴</sup> Action

- 
- <sup>°</sup> Unfavorable
  - <sup>˘</sup> Stationary
  - <sup>˘</sup> Non-Stationary
  - <sup>^</sup> fixed state transition probabilities
  - <sup>ˆ</sup> Variable Structure Learning Automata
  - <sup>ˆˆ</sup> Linear Reward Penalty
  - <sup>ˆˆ</sup> Linear Reward Epsilon Penalty
  - <sup>ˆˆ</sup> Linear Reward Inaction
  - <sup>ˆˆ</sup> Tutorial-like system
  - <sup>ˆˆ</sup> Synchronous
  - <sup>ˆˆ</sup> Sequential
  - <sup>ˆˆ</sup> Socratics
  - <sup>ˆˆ</sup> Fixed Structure Automaton
  - <sup>ˆˆ</sup> Variable Structure Automaton



# کنفرانس داده کاوی ایران