

# بهبود مدلسازی رفتار دانشجو در سیستم شبیه آموزشی با استفاده از آutomاتای

## یادگیر

لیلا جوادی<sup>۱</sup> ; بهروز معصومی<sup>۲</sup> ; محمد رضا 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 شبیه آموزشی با استفاده از آutomاتای یادگیر بیان می کند. در مدل پیشنهادی، مدل دانشجو توسط یک آtomاتای سطح بالاتر به نام Meta\_LA که مدل یادگیری دانشجو را مشخص می کند، تعیین می شود. Meta\_LA در حقیقت آtomاتای یادگیر را به عنوان مکانیزم یادگیری برای تعیین اینکه دانشجو یادگیرنده کند، نرم افزار یا سریع می باشد، بکار می گیرد. در این مقاله با به کارگیری آtomاتای یادگیر Pursuit دقت مدل یادگیری دانشجو افزایش یافته است.

### کلمات کلیدی

آtomاتای یادگیر، مدلسازی دانشجو، سیستم های شبیه 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\_javadi87@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 سیستم‌های رایانه‌ای در امور آموزشی هر روز بیشتر شده است. بیشتر سیستم‌های 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 می‌نمایند، 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 می‌کند که هر فرد سبک یادگیری مخصوص خود را دارد. بنابراین در یک سیستم آموزشی هوشمند با استفاده از تکنیک‌های هوش مصنوعی، براساس نیاز یادگیرنده، موضوع آموزشی به صورت اختصاصی برای وی آماده و در بهترین روش ارائه می‌گردد.

سامانه‌های آموزشی هوشمند، نسل جدیدی از سامانه‌های آموزشی هستند که شامل تکنیک‌های هوش مصنوعی و مندهای ارائه داشتند. بوده و به صورت واکنش پذیری بر اساس نیاز یادگیرنده توصیف شده که با هدف بهبود آموزش و یادگیری انسان‌ها بوجود آمده‌اند. این سیستم‌ها امکان تقلید سازی از یک معلم انسانی را بوجود آورده به طوریکه می‌توان آنچه را که تدریس می‌شود و همچنین چگونگی تدریس را کنترل نمود [۱ و ۲]. برای سیستمهای آموزشی هوشمند با توجه به تعابیر مختلف از مفهوم هوشمند سازی و نیز حوزه‌های کاربردی مختلف، معماری‌ها و اجزاء گوناگونی پیشنهاد شده است. آنچه در تمامی این طرح‌ها تقریباً مشترک است، مفهومی از تعامل بین یادگیرنده و مدل‌هایی از دانش دامنه مورد نظر، افراد متخصص و مدل استخراج شده از یادگیرنده می‌باشد. در حالت کلی این سیستم‌ها اساساً شامل سه مولفه اصلی هستند و گاهی مولفه واسطه کاربری نیز به آن اضافه می‌شود [۳ و ۴]. سه مولفه تشکیل دهنده سیستم آموزشی هوشمند شامل مدل دامنه (دانش دامنه)، مدل دانشجو و مدل آموزشی<sup>۱</sup> می‌باشند. Self این مولفه‌ها 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 سیستم‌های آموزشی در تعدادی از تحقیقات گذشته مورد مطالعه بوده است. Frasson در سال ۱۹۹۶ مولفه‌های اصلی سیستم آموزشی هوشمند (مدل دانشجو، دانش دامنه و مدل آموزشی) را به شکل عامل‌های هوشمند طراحی کرده است [۶]. Lelouche در سال ۲۰۰۰ مجموعه‌ای از عامل‌های تعاملی را برای مدل‌سازی اصلی دانش آموزشی در سیستم آموزشی هوشمند بکار می‌گیرد [۷]. 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 وضعیت‌های معین از دانش دانشجو استفاده کرده است [۹]. Mooney و Baffles در سال ۱۹۹۶ ASSERT را پیاده سازی کردن که یادگیری تقویتی را در مدل‌سازی دانشجو برای بدست آوردن خطاهای دانشجو جدید فقط با استفاده از دانش دامنه صحیح استفاده می‌کند [۱۰]. hashem در سال ۲۰۰۷ استفاده از آتماتای 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 دنباله‌ای از اعمال انجام داده دانشجو را ثبت کرده و با استفاده از این سیستم الگوی یادگیری بکار گرفته شده توسط دانشجو را مدل کرده و پیشرفت یادگیری دانشجو را شناسایی می‌کند.

هدف این مقاله بهبود مدل‌سازی دانشجو در سیستم یادگیری می‌باشد و نشان داده شده است که چگونه با استفاده از آتماتای یادگیر و در چارچوب سیستم شبه آموزشی، می‌تواند این مدل‌سازی موفقیت آمیز باشد. سیستم آموزشی برای بالابدن منحنی یادگیری دانشجو نیاز دارد تا با هر دانشجو بر اساس مهارت و قابلیت رفتار کند. به عبارتی دیگر این مقاله دیدگاه جدیدی را برای اینکه فرایند یادگیری دانشجو در سیستم شبه آموزشی مبتنی بر تئوری آتماتای یادگیر تصادفی مدل شود، بیان می‌کند. آتماتای یادگیر تصادفی سیستم را برای تشخیص مدل یادگیری دانشجو توانا می‌سازد و در نتیجه به سیستم شبه آموزشی جهت بهبود روش رفتار (مانند تدریس و ارزیابی) با دانشجو متناسب با مهارت او کمک می‌کند. برای بهبود تعیین مدل دانشجو با توجه به الگوریتم پیشنهادی به نام<sup>۲</sup> DP<sub>RI</sub> در مدل DL<sub>RI</sub> مرجع [۱۱] استفاده شده است که با توجه به آزمایش‌های مختلف نشان داده شده است مدل بهبود یافته با دقت بیشتری مدل دانشجو را تشخیص می‌دهد. در صورتی که استاد بتواند مدل یادگیری دانشجو را تشخیص دهد، می‌تواند استراتژی تدریس را با مدل یادگیری دانشجو منطبق کند. برای مثال اگر استاد بداند که با یک دانشجو زرنگ سروکار دارد، آنگاه مسائل با پیچیدگی و سرعت بیشتر را برای دانشجو ارائه می‌دهد. در مدل پیشنهادی فرض بر این است که دانشجو می‌تواند به سه صورت دانشجوی گند، دانشجوی نرمال و دانشجوی سریع مدل شود. سیستم، مدل یادگیری دانشجو را با بررسی اعمال متواالی دانشجو و با بکارگیری Meta-LA تعیین می‌کند [۱۱ و ۱۲].

برای ارزیابی مدل مطرح شده از محیط‌های مختلفی شامل محک‌های ثابت تست استفاده شد و نتایج بدست آمده از شبیه سازی نشان می‌دهند 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 آتماتای 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. آنوماتای یادگیر

آنوماتای یادگیر یکی از مدل‌های یادگیری تقویتی است که در آن آنوماتا یک عمل<sup>\*</sup> بهینه را از میان یک مجموعه از عمل‌های مجاز آنوماتا انتخاب می‌کند. این عمل روی یک محیط تصادفی اعمال می‌شود و محیط به این عمل آنوماتا بوسیله یک پاسخ تصادفی از مجموعه پاسخ‌های مجاز جواب می‌دهد. پاسخ محیط بصورت آماری به عمل آنوماتا وابسته است. اصطلاح محیط شامل تمام شرایط خارجی و تاثیرات آنها روی عملکرد آنوماتا است. ساختار یک آنوماتای یادگیر با محیط در شکل ۱ نشان داده است.



شکل ۱: اتصال آنوماتای یادگیر با محیط [۱۵]

محیط را می‌توان توسط سه تایی  $E = \{\alpha, \beta, c\}$  نشان داد که در آن  $\alpha = \{\alpha_1, \alpha_2, \dots, \alpha_r\}$  مجموعه ورودی‌ها،  $\beta = \{\beta_1, \beta_2, \dots, \beta_m\}$  مجموعه خروجی‌ها و  $c = \{c_1, c_2, \dots, c_r\}$  مجموعه احتمال‌های جریمه می‌باشد. هرگاه  $\beta$  مجموعه دو عضوی باشد، محیط از نوع P می‌باشد، در چنین محیطی  $\beta_1 = 1$  به عنوان جریمه و  $\beta_2 = 0$  به عنوان پاداش در نظر گرفته می‌شود. در محیط از نوع Q،  $\beta(n)$  می‌تواند به طور گسته‌یک 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  $i$  نتیجه نامطلوب<sup>۵</sup> داشته باشد، می‌باشد. در محیط ایستا<sup>۶</sup> مقادیر  $C_t$  بدون تغییر می‌مانند، حال آن که در محیط غیر ایستا<sup>۷</sup> این مقادیر در طی زمان تغییر می‌کنند. آنوماتای یادگیر به دو گروه با ساختار ثابت و با ساختار متغیر تقسیم می‌گردد<sup>[۱۳] و [۱۴]</sup> و <sup>[۱۵]</sup>. آنوماتای یادگیر با ساختار ثابت با احتمالات گذر وضعیت ثابت مشخص<sup>۸</sup> می‌شود. نظریه زنجیره‌های مارکوف ابزار اصلی تجزیه و تحلیل این کلاس از آنوماتا می‌باشد و در اغلب موارد، رفتار مقتضی با انتخاب احتمالات گذر وضعیت آنوماتا در پاسخ به خروجی محیط به دست می‌آید.

آنوماتای یادگیر با ساختار متغیر<sup>۹</sup> توسط ۴ تایی  $\{\alpha, \beta, p, T\}$  نشان داده می‌شود که در آن  $\alpha = \{\alpha_1, \alpha_2, \dots, \alpha_r\}$  مجموعه عمل‌های آنوماتا،  $\beta = \{\beta_1, \beta_2, \dots, \beta_m\}$  مجموعه ورودی‌های آنوماتا و  $p = \{p_1, p_2, \dots, p_n\}$  بردار احتمال انتخاب هر یک از اعمال و  $T[\alpha(n), \beta(n), p(n)] = p(n+1)$  الگوریتم یادگیری می‌باشد. در این نوع از آنوماتاهای، اگر عمل  $i$  در مرحله n ام انتخاب شود و این عمل، پاسخ مطلوب از محیط دریافت نماید، احتمال  $(n)_i$  افزایش یافته و سایر احتمال‌ها کاهش می‌یابند. برای پاسخ نامطلوب احتمال  $(n)_i$  کاهش یافته و سایر احتمال‌ها افزایش می‌یابند. در هر حال، تغییرات به گونه‌ای صورت می‌گیرد تا حاصل جمع  $(n)_i$  ها مساوی باقی بماند. الگوریتم زیر بر اساس (۲) و (۱) نمونه‌ای از الگوریتم‌های یادگیری خطی در آنوماتای با ساختار ثابت است<sup>[۱۵]</sup>.

الف- پاسخ مطلوب برای عمل  $i$ :

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

(۲)

در روابط فوق،  $a$  پارامتر پاداش و  $b$  پارامتر جریمه می‌باشد. با توجه به مقادیر  $a$  و  $b$  سه حالت زیر را می‌توان در نظر گرفت. زمانی که  $a$  و  $b$  با هم برابر باشند، الگوریتم را  $L_{RP}$ <sup>۱۰</sup> می‌نامیم. زمانیکه  $b$  از  $a$  خیلی کوچکتر باشد، الگوریتم را  $L_{REPL}$ <sup>۱۱</sup> می‌نامیم و زمانی که  $b$  مساوی صفر باشد، الگوریتم را  $L_{RI}$ <sup>۱۲</sup> می‌نامیم.

آنوماتاهای یادگیر دارای کاربردهای فراوانی می‌باشند. بعضی از این کاربردها عبارتند از: مسیریابی در شبکه‌های ارتباطی، فشرده سازی تصاویر، شناسایی الگو، زمانبندی فرآیندها در شبکه‌های کامپیوتری، تئوری صفر، کنترل دسترسی در شبکه‌های انتقال ناهمزمان، کمک به آموزش شبکه‌های عصبی، دسته بندی و افزار اشیاء و پیدا کردن ساختار بهینه برای شبکه‌های عصبی<sup>[۱۶] و [۱۷]</sup>.

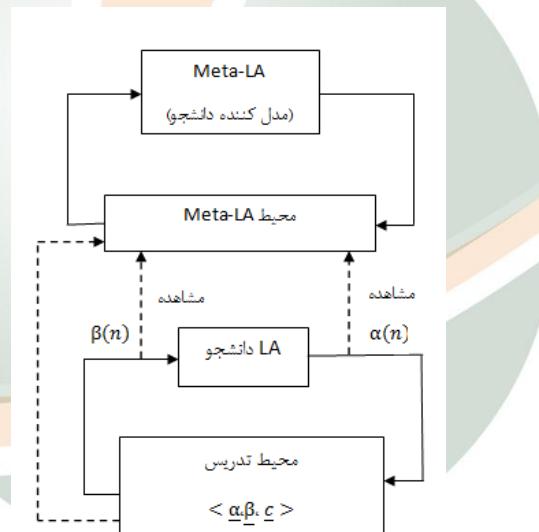
### ۳. سیستم شبه آموزشی<sup>۱۲</sup>

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 شبیه 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 شبیه ساز دانشجویی که آتماتای 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 آتماتای یادگیری اجازه خواهد داد مدل شبیه ساز خصوصیاتی از دانشجوی واقعی را تشخیص داده و به همان اندازه تعاملات دانشجو با سیستم شبه آموزشی را افزایش دهد و همچنین یادگیری از آتماتای یادگیری افزایش می یابد، که بر این دلالت دارد که مدل شبیه ساز، یادگیری بهتری از دانشجو ارائه می کند. این انتظار وجود دارد که بعد از تکمیل سیستم، مدل شبیه ساز دانشجو می تواند توسط دانشجویان واقعی جایگزین شود [۱۲].

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

سیستم شبه آموزشی از چند آتماتای یادگیر تشکیل شده است که بطور غیر مستقیم به همدیگر وصل شده اند (یک سیستم آتماتای بهم متصل است). نحوه اتصال آتماتاهای 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 آتماتای یادگیر در این سیستم در شکل ۲ نشان داده شده است [۱۱ و ۱۶ و ۲۰].



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

در آتماتای متصل سنتی پاسخ محیط یک آتماتا، ورودی آتماتای دیگر را بیان می کند [۱۵]، در حالی که در این سیستم، تا زمانیکه آتماتای یادگیر دانشجویان روی Meta-LA تاثیر گذار باشد، مابین آنها ارتباط مستقیمی وجود ندارد. محیط Meta-LA کارایی آتماتای 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 آتماتای سطح بالا مثل Meta-LA و یک آتماتای سطح پایین که آتماتای یادگیر دانشجو است. همگرایی آتماتای سطح بالا به رفتار آتماتای سطح پایین بستگی دارد. یک سناریو نمونه می گوید که آتماتای سطح بالا می تواند قبل از همگرایی آتماتای سطح پایین، همگرا شود در حالیکه اگر آتماتای سطح پایین قبل از آتماتای سطح بالا همگرا شود، این به معنای همگرایی نادرست آتماتای سطح بالا است.

از مشاهده تعامل ماژول 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 آتماتای یادگیر سطح پایین ناشناخته است، دسترسی دارد، Meta-LA به اعمال انتخاب شده توسط آتماتای یادگیر سطح پایین و نیز پاسخ های ایجادشده 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 ۱۶ و ۲۰].

## ۴. مدل پیشنهادی و بهبود مدلسازی دانشجو

سیستم شبیه آموزشی یک مدل را برای دانشجو توسط ارزیابی و تعیین راهی که او یاد می‌گیرد، استنتاج می‌کند. برای رسیدن به این، فرض شده که سیستم یک مجموعه متناهی از مدل‌های 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 می‌تواند کاملاً یا بطور جزئی برای وجههای خاصی از رفتار دانشجو محاسبه شود. دامنه دانش با سوالات مدل سقراطی<sup>۱۶</sup> چند گزینه‌ای نمایش داده می‌شود 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 می‌تواند با یکی از سه نوع، یک مدل برگرفته از آtomاتای یادگیر با ساختار ثابت که آن را<sup>۱۷</sup> FSSA می‌نامیم برای یک دانشجوی کند، مدل مبتنی بر آtomاتای یادگیر متغیر<sup>۱۸</sup> VSSA برای دانشجوی نرمال و مدل مبتنی بر آtomاتای Pursuit برای دانشجوی سریع مدل شود. این مجموعه متناهی از مدل‌های یادگیری راههای مختلف یادگیری دانشجو را ارائه می‌دهد. اگر سیستم شبیه آموزشی بتواند چگونگی درک دانشجو را بفهمد، آنگاه می‌تواند استراتژی تدریس بهینه را برای هر دانشجو پیدا کند. برای مثال، اگر سیستم شبیه آموزشی تشخیص دهد که دانشجوی در حال یادگیری، مشابه مدل Pursuit است، آنگاه می‌تواند نتیجه بگیرد که دانشجو زرنگ است و اینکه آن یک یادگیرنده سریع می‌باشد. در این حالت سیستم می‌تواند استراتژی تدریس را برای آمادگی یک دانشجوی سریع انطباق دهد که آن را برای بهبود استراتژی تدریس توana می‌سازد.

مدل دانشجو از سه نوع خاص آtomاتای یادگیری استفاده خواهد کرد، هریک از اینها به صورت یک عمل برای آtomاتای سطح بالا مورد استفاده قرار می‌گیرند. مدل آtomاتای یادگیر یک مدل با شماهای پیوسته است که، در این حالت احتمالات عمل می‌توانند هر مقدار واقعی در بازه [۰,۱] باشند. Oommen و Thathachar مفهوم آtomاتای یادگیر گستره را توسط محدود کردن احتمال انتخاب از یک مجموعه متناهی از مقادیر در بازه بسته [۰,۱] مطرح کردند. بنابراین احتمال عمل در گام‌ها به صورت گستره بروزرسانی می‌شود به جای آنکه بروزرسانی در حالت پیوسته باشد [۲۲]. مهمترین مزیت استفاده از الگوریتم‌های گستره افزایش سرعت همگرایی است [۲۳]. مزیت دیگر استفاده از این الگوریتم‌ها حذف شدت اعمال نفوذ بر روی مولد اعداد تصادفی است و مزیت سوم آن، استفاده کمتر از حافظه و اجرای سریعتر از آtomاتای پیوسته است. Oommen اثبات کرد که آtomاتاهای گستره در همه محیط‌ها نزدیک به بهینه بوده و از همتاها پیوسته شان همگرایی سریعتری دارند [۲۴]. در روش پیشنهادی از الگوریتم‌های تعقیبی استفاده شده است که درواقع، کلاسی از الگوریتم‌های تخمین زننده است و عمل بهینه را پیگیری می‌کند و شامل سه مرحله است، مرحله اول عملی بر مبنای توزیع احتمال P(t) انتخاب شده و در مرحله دوم اگر آtomاتا پاداش بگیرد، مولفه ای از p(t) که تخمین پاداش جاری اش ماکزیمم است افزایش می‌یابد و احتمالات عمل بقیه کاهش می‌یابد و نهایتاً در مرحله سوم تخمین های جاری احتمال بروزرسانی می‌شود. این الگوریتم را DP<sub>RI</sub> می‌نامیم. به نظر می‌رسد الگوریتم DP<sub>RI</sub> باید همگرایی سریعتری نسبت به DL<sub>RI</sub> داشته باشد و در نهایت مدلسازی دانشجو بهبود داده می‌شود.

در مثال ارائه شده Meta-LA برای یادگرفتن بهترین مدل ممکن برای شبیه سازی دانشجو از چهارتایی  $\{\alpha, \beta, P, T\}$  استفاده می‌کند که در آن مجموعه اعمال به صورت  $\{\alpha_1, \alpha_2, \alpha_3\} = \underline{\alpha}$  است.  $\alpha_1$  نشان می‌دهد عمل با FSSA، یادگیرنده کند مطابق است،  $\alpha_2$  نشان می‌دهد عمل با VSSA، یادگیرنده نرمال مطابق است و  $\alpha_3$  نشان می‌دهد عمل با Pursuit، یادگیرنده سریع مطابق است. اشاره می‌کند که پاداش برای عمل حاضر انتخاب شده است و  $\beta = 0,1$  اشاره می‌کند که در نهایت انتخاب شده است و  $T$  قانون بروزرسانی احتمالات است که بصورت مقابل نگاشت می‌شود:  $P \rightarrow (P, \alpha, \beta)$ . هریک از اینها در ادامه بصورت کامل توضیح داده شده اند:

برای تعداد ثابت پرس و جوها، Meta-LA فرض کند مدل شبیه ساز دانشجو  $\alpha_i$  باشد و  $p_i(n) = [p_1, p_2, p_3]^T$  احتمالات جاری شبیه ساز دانشجو که با  $\alpha_i$  نشان داده شده است و  $T$  نرخ یادگیری آtomاتای یادگیرسطح پایین در زمان t باشد. توسط بازبینی راهی که شبیه ساز دانشجو در مدت این پرس و جوها یاد می‌گیرد، استنتاج می‌شود که آیا مدل جاری باید پاداش بگیرد یا جرمیه شود؟ ما توسط دو آستانه به هدف مان می‌رسیم:  $\{j \leq 1 | \theta_j \leq \theta\}$  اگر نرخ یادگیری مشاهده شده جاری از شبیه ساز دانشجو کمتر از  $\theta_1$  باشد، ادعا می‌کنیم که شبیه ساز دانشجو یادگیرنده سریع است. حال اگر Meta-LA را انتخاب کند، سیستم پاداش را انتخاب می‌کند در غیر اینصورت جرمیه انتخاب می‌شود.

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

خصوصیات یادگیرنده کند را دارد. حال اگر  $\alpha_3$  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  $\text{Meta-LA}$  به هر یک از اعمالش نسبت می دهد. در لحظه  $t$   $\text{Meta-LA}$  بطور تصادفی یک عمل  $i$  را انتخاب می کند. احتمالی که آتماتا عمل  $i$  را در زمان  $t$  انتخاب کند، احتمال عمل است و در هر حال، تغییرات به گونه ای صورت می گیرد تا حاصل جمع  $(n)$  ها همواره ثابت و مساوی یک باقی بماند که در (۳) نشان داده شده است.

جدول ۱ آستانه مورد استفاده توسط  $\text{Meta-LA}$  برای مدل یادگیری [۱۱]

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

$$p_i(t) = pr[\alpha(t) = \alpha_i], \quad (3)$$

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

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

$$p(0) = [0.333, 0.333, 0.333]^T \quad (4)$$

ت، الگوریتم بروز رسانی یا شما کمکی است که برای بروز رسانی بردار احتمال عمل  $\text{Meta-LA}$  بکار گرفته می شود و بصورت (۵) نمایش داده می شود:

$$p(t+1) = T[p(t), \alpha(t), \beta(t)] \quad (5)$$

این الگوریتم بروز رسانی می تواند هر یک از شماهای یادگیری  $\text{LA}$  باشد. در روش پیشنهادی از شما  $\text{DP}_{\text{RI}}$  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 بروزرسانی نمی شود [۲۵]. شما  $\text{DP}_{\text{RI}}$  در (۶) بیان شده است.

(۶)

Else

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

به عنوان یک آتماتای یادگیر سطح بالا پیاده سازی شده است و همانطور که در قسمت ۱-۳ بیان شده است به محیط یادگیری، اعمال و جریمه های آتماتای یادگیر سطح پایین وابسته به شبیه ساز دانشجو دسترسی دارد بطوریکه بعدا مواد تدریسی را یاد می گیرد.  $\text{Meta-LA}$  به عنوان یک آتماتای یادگیر گسته با شما  $\text{DP}_{\text{RI}}$  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 استفاده می کند. در اصل از آتماتای  $\text{Tsetlin}$  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  $\text{VSSA}$  استفاده می کند که در این مقاله با شما  $L_{\text{RI}}$  و با  $\lambda = 0.005$  پیاده سازی شده است.
- یادگیرنده سریع: برای شبیه سازی دانشجو از این نوع، شبیه ساز دانشجو از آتماتای یادگیر  $\text{Pursiut}$  استفاده می کند که در این مقاله با شما  $PL_{\text{RI}}$  و با  $\lambda = 0.005$  پیاده سازی شده است.



جدول ۳ مقایسه همگرایی Meta\_LA در مدل پیشنهادی و مدل پیشنهادشده درمرجع [۱۱] در محیط E(B,f)

محیط E(B,f)			PL <sub>RI</sub>		L <sub>RI</sub>		Tsetlin	
			تعداد تکرار	همگرایی صحیح	تعداد تکرار	همگرایی صحیح	تعداد تکرار	همگرایی صحیح
مدل پیشنهادشده درمرجع [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 <sub>B</sub> = { ۰,۱ ۰,۴۵ ۰,۸۴ 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 آtomاتای سطح پایین تراز 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 درمرجع [۱۱]، %۹۲ می باشد. همچنین در زمان استفاده از آtomاتای L<sub>RI</sub> درصد همگرایی صحیح در الگوریتم ما %۱۰۰ و در الگوریتم پیشنهادشده درمرجع [۱۱]، %۱۰۰ بوده است که در این حالت الگوریتم ما با تعداد تکرار کمتری به این نتیجه رسیده است و همینطور در زمان استفاده آtomاتای سطح پایین تراز 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 موقع موفق بوده و در تعیین دانشجوی سریع در %۹۴

جدول ۴ مقایسه همگرایی Meta\_LA در مدل پیشنهادی و مدل پیشنهادشده درمرجع [۱۱] در محیط E(A,10)

محیط E(A,10)			PL <sub>RI</sub>		L <sub>RI</sub>		Tsetlin	
			تعداد تکرار	همگرایی صحیح	تعداد تکرار	همگرایی صحیح	تعداد تکرار	همگرایی صحیح
مدل پیشنهادشده درمرجع [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 ۰,۰۴ ۰,۰۵ ۰,۰۴ ۰,۰۳ ۰,۰۵ ۰,۰۲ }					

جدول ۵ مقایسه همگرایی Meta\_LA در مدل پیشنهادی و مدل پیشنهادشده درمرجع [۱۱] در محیط E(B,10)

محیط E(B,10)			PL <sub>RI</sub>		L <sub>RI</sub>		Tsetlin	
			تعداد تکرار	همگرایی صحیح	تعداد تکرار	همگرایی صحیح	تعداد تکرار	همگرایی صحیح
مدل پیشنهادشده درمرجع [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 ۰,۰۲ ۰,۰۴ ۰,۰۳ ۰,۰۵ ۰,۰۳ }					

### 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 آtomاتای یادگیری 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 آtomاتای سطح بالاتر به نام Meta-LA استفاده می کند که اعمال آtomاتاهای شبیه ساز دانشجو و محیط تدریس را مشاهده می کند و تلاش می کند تا مدل یادگیری دانشجو را تعیین کند. مدل پیشنهادشده درمرجع [۱۱] 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 حالت‌ها، 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, ۱۹۹۰.
- [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 ۵۷-۶۰. Springer-verlag, Berlin, ۱۹۹۶.
- [۷] R.Lelouche; “*A collection of pedagogical agents for intelligent educational systems*”, In intelligent tutoring systems: ۷<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. ۱۰th Nat. Conf. AI/IAAI, Madison, WI, p. ۱۱۸۰, ۱۹۹۸.
- [۱۰] P. Baffes ; R. Mooney; “*Refinement-based student modeling and automated bug library construction*”, J. Artif. Intell. Educ, vol. ۴, no. ۱, pp. ۷۵-۱۱۶, ۱۹۹۷.
- [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, . . . © 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/1 Queuing System with Unknown Parameters*”, Proceedings of Third Yale Workshop on Applications of Adaptive System Theory, Yale University, ۱۰-۱۰-۹, ۱۹۸۳.
- [۱۷] M. R. Meybodi ; H. Beigy; “*New Class of Learning Automata Based Scheme for Adaptation of Backpropagation Algorithm Parameters*”, Proceedings of EUFIT-۹۸, Achen, 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. ۴۰۷۰, 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

- 
- Unfavorable
  - Stationary
  - Non-Stationary
  - ^ fixed state transition probabilities
  - ^ Variable Structure Learning Automata
  - Linear Reward Penalty
  - Linear Reward Epsilon Penalty
  - Linear Reward Inaction
  - Tutorial-like system
  - Synchronous
  - Sequential
  - Socratics
  - Fixed Structure Automaton
  - Variable Structure Automaton



# کنفرانس داده کاوی ایران