

# بهبود مدلسازی دانشجو در سیستم شبه آموزشی با استفاده از یادگیری

## تقویتی و آتماتای یادگیر

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

<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.

mneybodi@aut.ac.ir

### چکیده

در جامعه کنونی به کارگیری کامپیوتر جهت آموزش امری ضروری است و در این میان سیستم های آموزشی هوشمند توجه محققان را به خود جلب کرده است. مدلسازی دانشجو نقطه تمرکز اصلی و یکی از اجزاء سیستم آموزشی هوشمند و 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 سیستم شبه آموزشی با استفاده از آتماتای یادگیر و یادگیری تقویتی بیان می کند. در مدل پیشنهادی، مدل دانشجو توسط الگوریتم یادگیری تقویتی سطح بالاتر به نام Meta Q/LA که مدل یادگیری دانشجو را مشخص می کند، تعیین می شود. Meta Q/LA در حقیقت الگوریتم یادگیری Q را به عنوان مکانیزم یادگیری برای تعیین اینکه دانشجو یادگیرنده کند، نرم‌الی یا سریع می باشد، بکار می گیرد و نشان داده شد با بکارگیری الگوریتم یادگیری Q سرعت تعیین مدل یادگیری دانشجو افزایش یافته است.

### کلمات کلیدی

آتماتای یادگیر، مدلسازی دانشجو، یادگیری Q، سیستم های شبه آموزشی

### ۱ - مقدمه

معماری ها و اجزاء گوناگونی پیشنهاد شده است. در حالت کلی این سیستم ها شامل سه مولفه اصلی هستند و گاهی مولفه واسط کاربری نیز به آن اضافه می شود [۴،۳]. سه مولفه تشکیل دهنده سیستم آموزشی هوشمند شامل مدل دامنه، مدل دانشجو و مدل آموزشی<sup>۱</sup> می باشد [۵].

به کارگیری یادگیری ماشین در بهبود سیستم های آموزشی در تعدادی از تحقیقات گذشته مورد مطالعه قرار گرفته است. Lelouch در سال ۲۰۰۰ مجموعه ای از عامل های تعاملی را برای مدلسازی اصلی دانش، آموزشی در سیستم آموزشی هوشمند بکار گرفت [۶]. Legaspi و Sison در سال ۲۰۰۰ آموزش را در سیستم آموزشی هوشمند با استفاده از یادگیری تقویتی به عنوان رویه یادگیری مرکزی مدل کردند [۷]. Mooney و Baffles در سال ۱۹۹۶ ASSERT را پیاده سازی کردند که یادگیری تقویتی را در مدلسازی دانشجو برای

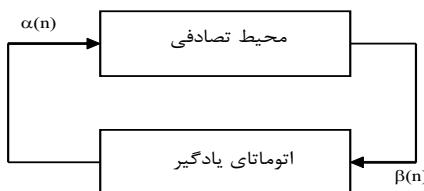
در طول بیست سال اخیر حضور گسترده سیستم های کامپیوتری در امور آموزشی هر روز بیشتر شده است. سازگاری و تطابق در آموزش، به منظور افزایش کارایی و تاثیر بر فرایند یادگیری ضروری است. اما تعداد کمی از سیستم های آموزشی وجود دارند که پویا بوده و قادراند مطالب آموزشی را به صورت شخصی متناسب با نیازها و سبک یادگیری هر دانش آموز فراهم نمایند [۱]. برنامه هایی را که با استفاده از تکنیک های هوش مصنوعی به افراد در یادگیری کمک می کنند، "سیستم های آموزشی هوشمند" می نامند.

سامانه های آموزشی هوشمند، نسل جدیدی از سامانه های آموزشی هستند که شامل تکنیک های هوش مصنوعی و متدهای ارائه دانش بوده که با هدف بهبود آموزش و یادگیری انسان ها بوجود آمدند [۲،۱]. برای سیستم های آموزشی هوشمند با توجه به تعابیر مختلف از مفهوم هوشمند سازی و نیز حوزه های کاربردی مختلف،

این الگوریتم تعیین می کند که به ازای مقادیر  $Q$  صحیح اگر محیط ثابت و وابسته به محیط جاری و عمل انجام شده در آن باشد، با احتمال یک همگرا می شود [۱۱].

### ۳- آتماتای یادگیر

آتماتای یادگیر یکی از مدل های یادگیری تقویتی است که در آن آتماتای یک عمل بهینه را از میان یک مجموعه از عمل های مجاز آتماتا انتخاب می کند. این عمل روی یک محیط تصادفی اعمال می شود و محیط به این عمل آتماتا بوسیله یک پاسخ تصادفی از مجموعه پاسخ های مجاز جواب می دهد. پاسخ محیط بصورت آماری به عمل آتماتا وابسته است. اصطلاح محیط شامل تمام شرایط خارجی و تاثیرات آنها روی عملکرد آتماتا است. ساختار یک آتماتای یادگیر با محیط در شکل (۱) نشان داده شد.



شکل (۱) اتصال آتماتای یادگیر با محیط [۱۲]

آتماتای یادگیر به دو گروه با ساختار ثابت و با ساختار متغیر تقسیم می شود [۱۲، ۱۴، ۱۳]. آتماتای یادگیر با ساختار ثابت با احتمالات انتقال ثابت مشخص<sup>۴</sup> می شود. آتماتای یادگیر با ساختار متغیر<sup>۵</sup> توسط ۴ تایی  $\{\alpha, \beta, p, T\}$  نشان داده می شود، که در آن  $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$  مجموعه اعمال آتماتا،  $\beta \equiv \{\beta_1, \beta_2, \dots, \beta_m\}$  مجموعه ورودی های آتماتا 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 بروزرسانی است. در این نوع از آتماتاها، اگر عمل  $i$  در مرحله  $n$  انتخاب شود و این عمل، پاسخ مطلوب از محیط دریافت کند، احتمال  $p_i$  افزایش یافته و سایر احتمال ها کاهش می یابد. برای پاسخ نامطلوب احتمال  $(n)_i$  کاهش یافته و سایر احتمال ها افزایش می یابد. در هر حال، تغییرات به گونه ای صورت می گیرد تا حاصل جمع  $(n)_i$  ها همواره ثابت و مساوی یک باقی بماند.

### ۴- سیستم شبه آموزشی<sup>۶</sup>

سیستم های شبه آموزشی، نمایشی از وضعیت دانشجو می باشد. این سیستم، توانایی کارکردن و تست بدون نیاز به اشخاص واقعی را دارد. 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

بدست آوردن خطاهای دانشجو جدید فقط با استفاده از دانش دامنه صحیح استفاده کردد [۸]. hashem در سال ۲۰۰۷ استفاده از آتماتای یادگیر را در سیستم آموزشی هوشمند بکار گرفت [۹].

مدلسازی دانشجو نقطه تمرکز اصلی و یکی از اجزاء سیستم آموزشی است. این مدل، نمایشی از رفتار و وضعیت دانشجو بوده و مبنای برای نمایش وضعیت دانشجو است. هدف این مقاله بهبود مدلسازی دانشجو در سیستم یادگیری می باشد و نشان داده شد که چگونه با استفاده از الگوریتم یادگیری  $Q$  و در چارچوب سیستم شبه آموزشی، این مدلسازی می تواند موقیت آمیز باشد. برای بهبود تعیین مدل دانشجو از الگوریتم یادگیری  $Q$  به جای الگوریتم  $DL_{RI}$  در مدل مرجع [۹] استفاده شد که با توجه به آزمایش های مختلف نشان داده شد که مدل بهبود یافته با سرعت بیشتری مدل دانشجو را تشخیص می دهد [۹، ۱۰].

برای ارزیابی مدل مطرح شده از محیط های مختلفی شامل مک های ثابت تست استفاده شد و نتایج بدست آمده از شبیه سازی نشان داد که Meta Q/LA در تعیین مدل یادگیری بر اساس مشاهده اعمال دانشجو و محیط تدریس موفق است. در ادامه مقاله در بخش ۲ به مفهوم یادگیری  $Q$  و در بخش ۳ مفهوم آتماتای یادگیری و انواع آن پرداخته شد. در بخش ۴ سیستم شبه آموزشی توضیح داده شد و در بخش ۵ مدل پیشنهادی بهبود یافته ارائه شد. در بخش ۶ به ارزیابی مدل پیشنهادی پرداخته و در بخش ۷ نتیجه گیری بیان شد.

### ۲- یادگیری $Q$

یادگیری  $Q$  یک شکل از الگوریتم یادگیری تقویتی است که نیاز به مدل محیط ندارد و می تواند به صورت آنلاین استفاده شود. بنابراین برای بازی های تکرار پذیر در مقابل یک حریف ناشناخته بسیار مناسب است. الگوریتم های یادگیری  $Q$  به وسیله تخمین زدن ارزش جفت های (وضعیت ، عمل) کار می کند. ارزش  $Q(s, a)$  برای بدست آوردن مجموعه ای discounted از بازده آینده تعریف شد تا بر اساس انجام عمل  $a$  از موقعیت  $s$  و دنبال کردن یک سیاست بهینه، عمل بعدی را پیش بینی کند. این ارزش ها به یکباره یاد گرفته می شود، عمل بهینه در هر موقعیت زمانی. بعد از اینکه تعدادی عمل تصادفی در ابتدا معین شد ، مقادیر  $Q$  بر اساس تجربیات زیر تخمین زده می شود:

۱. از موقعیت جاری  $s$  عمل  $a$  را انتخاب کن. این سبب دریافت نتیجه فوری ۲ و رسیدن به موقعیت بعدی ۳ خواهد شد.
  ۲. بروزرسانی کردن  $(s, a) Q$  بر اساس تجربیات معادله(۱):
- $$Q(s, a) = R(s, a) + \gamma \max Q(s, a) \quad (1)$$
۳. برو به مرحله ۱

از مشاهده تعامل ماجول ها در شکل نتایجی که در ادامه بیان شده است استنتاج می شود: محیط Meta Q/LA ترکیبی از مشاهدات است، که باید بطور ضمنی بعد از انجام پردازش پیام های مشاهده شده، دریافت شود، Meta Q/LA به محیط سطح پایین که شامل مجموعه ای از احتمالات جرمیه می باشد و برای آtomاتای یادگیر سطح پایین ناشناخته است، دسترسی دارد، و همچنین به اعمال انتخاب شده توسط آtomاتای یادگیر سطح پایین دارد. در مجموع، همانطور که در شکل مشخص می باشد، این نتایج مجموعاً محیط Meta Q/LA را تشکیل می دهد.

## ۵- مدل پیشنهادی و بهبود مدلسازی دانشجو

سیستم شبیه 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>۷</sup> چند گزینه ای نمایش داد شد و گزینه با در هر سوال، به هر گزینه احتمال پاداشی وابسته شد و گزینه بیشترین احتمال پاداش، پاسخ سوال در نظر گرفته شد [۱۷].

دانشجویان را در سه سطح کند، نرمال و سریع در نظر گرفتیم، یک مدل برگرفته از آtomاتای یادگیر با ساختار متغیر به نام  $L_{RI}$  و با نرخ یادگیری  $0.025$  برای یک دانشجوی کند، مدل مبتنی بر آtomاتای یادگیر متغیر  $L_{RI}$  و با نرخ یادگیری  $0.005$  برای دانشجوی نرمال و مدل مبتنی بر آtomاتای Pursuit به نام  $PL_{RI}$  برای دانشجوی سریع در نظر گرفته شد. این مجموعه متناهی از مدل های یادگیری راه های مختلف یادگیری دانشجو را ارائه می دهد.

در Meta Q/LA برای یادگرفتن بهترین مدل ممکن برای شبیه سازی دانشجو از یادگیری  $Q$  بدون وضعیت استفاده شده است که در آن مجموعه اعمال به صورت  $\alpha = \{\alpha_1, \alpha_2, \alpha_3\}$  است.  $\alpha_1$  نشان دهنده مطابقت عمل با یادگیرنده کند است،  $\alpha_2$  نشان می دهد عمل با یادگیرنده نرمال مطابق است و  $\alpha_3$  نشان می دهد عمل با یادگیرنده سریع مطابق است.

$\beta = 0$ ،  $\beta = 0.1$  اشاره می کند که پاداش برای عمل حاضر انتخاب شده و  $\beta = 1$  اشاره می کند که جرمیه برای عمل حاضر انتخاب شده است.

$$Q_i(n) = [q_1, q_2, q_3]^T$$

شبیه ساز دانشجویی که با  $\alpha_i$  نشان داده شده، می باشد. هریک از اینها در ادامه بصورت کامل توضیح داده شده است.

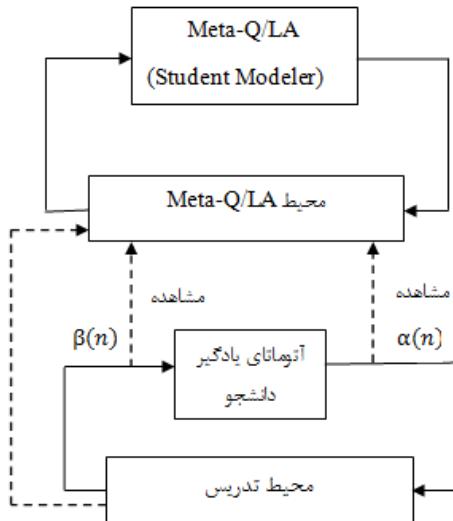
آموزشی را برای رسیدن به استراتژی بهینه بدون نیاز به اشخاص برای تعامل با یک سیستم آموزشی بهینه، توانا می سازد [۱۰].

در اینجا از شبیه ساز دانشجویی که الگوریتم یادگیری  $Q$  را بعنوان مکانیزم یادگیری برای تقلید چگونگی رفتار دانشجو نشان می دهد، استفاده شد. با به کارگیری یادگیری  $Q$  مدل شبیه ساز می تواند خصوصیاتی از دانشجو را تشخیص و به همان اندازه تعاملات دانشجو با سیستم شبیه آموزشی را افزایش دهد [۱۰].

## ۴-۱- ساختار شبکه آtomاتای یادگیر Meta Q/LA

سیستم شبیه آموزشی از چند آtomاتای یادگیر و یادگیری  $Q$  تشکیل شده است که بطور غیر مستقیم به یکدیگر وصل شدند. چگونگی اتصالات آtomاتای یادگیر و عامل  $Q$  در این سیستم در شکل ۲ نشان داده شد [۱۰، ۱۵، ۱۶].

در آtomاتای متصل سنتی پاسخ محیط یک آtomاتا، ورودی آtomاتای دیگر بود [۱۲]، در حالی که در این سیستم، تأثیر گذار باشد، آtomاتای یادگیر دانشجویان روی Meta Q/LA تأثیر گذار باشد، مابین آنها ارتباط مستقیم وجود ندارد. محیط Meta Q/LA کارایی آtomاتای یادگیر دانشجو را در یک دوره زمانی کنترل کرده و اینکه به عمل انتخابی، پاداش یا جرمیه دهد، به محیط آن بستگی دارد.



شکل (۲) مدل سازی با استفاده از شبکه ای از آtomاتاهای یادگیر این مدل یک ساختار جدید از ارتباطات داخلی را نشان می دهد که ترکیبی از دو سطح است. یک عامل  $Q$  مثل Meta Q/LA و یک آtomاتای سطح پایین که آtomاتای یادگیر دانشجو است. همگرایی عامل سطح بالا به رفتار آtomاتای سطح پایین بستگی دارد.

که در این معادله  $EV(action) = Q(action)$  در نظر گرفته شد. Meta Q/LA پاداش  $r$  را دریافت می کند و سپس براساس نرخ یادگیری  $\lambda$  که مقداری بین صفر و یک دارد و بر اساس (۳) بروزرسانی می شود.

$$Q_{t+1}(a) \leftarrow Q_t(a) + \lambda(r - Q_t(a)) \quad (3)$$

این عمل تا زمانیکه اختلاف  $Q_t(a)$  و  $Q_{t+1}(a)$  به حد آستانه  $T$  برسد ادامه می یابد. با این روش مدل یادگیری دانشجو تعیین می شود.

## ۶- ارزیابی مدل

در این قسمت نتایج بدست آمده از شبیه سازی مدل پیشنهادی نشان داده شد. برای بدست آوردن این نتایج، شبیه سازی های مختلف را اجرا کردیم تا مشخص کنیم مدل کننده دانشجو تا چه اندازه می تواند مدل یادگیری دانشجو را درست تشخیص دهد و نتایج بدست آمده با نتایج مدل قبلی [۹] مقایسه شد. شبیه سازی ها در انواع مختلفی از محیط ها تست شد که این محیط ها مسئله تدریس را نشان می دهد. در همه آزمایش های اجرا شده، اگر احتمال انتخاب یک عمل بزرگتر یا مساوی به آستانه  $T$  شود، الگوریتم همگرا در نظر گرفته می شود. اگر آtomاتا و یادگیری  $Q$  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، برای آtomاتا یادگیر  $0.99$  و برای یادگیری  $0.0001$  تنظیم شد و ۷۵ بار آزمایشات اجرا شد و نتیجه، میانگین ۲۵ 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 ۳۰۰ تکرار اول توسط Meta Q/LA صرفنظر شد. نتایج این شبیه سازی ها در ادامه توضیح داده شد.

### الف) آزمایش در محیط با ۱۰ عمل

محیط با ۱۰ عمل، یک سوال با ۱۰ گزینه را نمایش می دهد. احتمالات پاداش محیط و نتایج بدست آمده در محیط EA که بهترین عمل  $\alpha_1$  می باشد در جدول ۲ نشان داده شد. در زمانیکه آtomاتا سطح پایین از  $PL_{RI}$  استفاده می کند، در الگوریتم پیشنهادی، مدل یادگیری دانشجو را در ۹۶٪ زمان و با ۱۳۶ تکرار درست تشخیص داده است، در حالیکه این مقدار در الگوریتم پیشنهادشده درمرجع [۹] و ۴۶۸ تکرار می باشد. همچنین در تعیین دانشجوی نرمال، با میانگین ۱۴۸ تکرار و با دقت ۹۵٪ به همگرایی صحیح رسیده است، که این فاکتورها در الگوریتم پیشنهادشده درمرجع [۹] و ۵۴۸٪ بود. که در این حالت الگوریتم ما با تعداد تکرار کمتری به این

$\beta$  ورودی دریافتی Meta Q/LA است. دنباله ای از تصمیمات اخذ شده توسط شبیه ساز دانشجو را مشاهده کند و باید بر مبنای این دنباله نتیجه گیری کند که آیا مدل جاری اش درست است یا نه؟ برای رسیدن به این هدف پیشنهاد شده که، برای تعداد ثابت پرس و جوها، Meta Q/LA فرض کند مدل شبیه ساز دانشجو  $\alpha_i$  است. فرض می کنیم که  $\theta(t)$  نرخ یادگیری آtomاتا یادگیری 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؟ توسط دو آستانه به هدف مان می رسمیم:  $j \leq 2 | \theta_j \leq \theta_1$  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  $\alpha_1$  را انتخاب کند، سیستم پاداش را انتخاب می کند در غیر اینصورت جریمه انتخاب می شود.

جدول (۱) آستانه مورد استفاده توسط Meta Q/LA برای مدل

[۹] یادگیری

حد آستانه	انواع دانشجو
$\theta(t) < \theta_1$	یادگیرنده سریع
$\theta_1 \leq \theta(t) < \theta_2$	یادگیرنده نرمال
$\theta_2 \leq \theta(t)$	یادگیرنده کند

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  $\alpha_2$  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، می توان گفت 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.  $Q(n)$  بردار تخمین عمل است که شامل ارزشی است که Meta Q/LA به هر یک از اعمالش نسبت می دهد.

برای انتخاب عمل دو حالت جستجوگرانه<sup>۸</sup> و استناد بر پایه تجربیات<sup>۹</sup> مطرح است. در این مدل در لحظه  $t$  بر Meta Q/LA اساس توزیع بولتزمن که یک از روش های حالت جستجوگرانه استنادی است، عمل  $\alpha_i$  را بر اساس معادله (۲) انتخاب می کند.

$$P_i(action) = \frac{e^{\frac{EV(action)}{T}}}{\sum_{action \in A} e^{\frac{EV(action)}{T}}} \quad (2)$$

یادگیری Q با تعداد تکرار کمتری به دقت بالاتری بررسیم و این ساختار در تعیین مدل دانشجو موفق بود.

نتیجه رسیده است و همینطور در زمان تعیین دانشجوی کند این مقادیر به ترتیب ۱۵۹ تکرار و ۹۳٪ در الگوریتم پیشنهادی و ۱۱۶۹ تکرار و ۹۶٪ در الگوریتم پیشنهادشده در مرجع [۹] می باشد. همانطور که در جدول مشخص شد توانستیم با استفاده از

جدول (۲) مقایسه همگرایی Meta LA و Meta Q/LA در محیط E(A,10)

محیط E(A,10)	$\theta_1$	$\theta_2$	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 [۹]	۰,۴۶ - ۰,۳۵	۰,۵۷ - ۰,۴۰	۴۶۸	۹۵٪	۵۴۸	۹۳٪	۱۱۶۹	۹۶٪
مدل پیشنهادی			۱۳۶	۹۶٪	۱۴۸	۹۵٪	۱۵۹	۹۳٪
احتمال پاداش	$E_A = \{0.7, 0.5, 0.3, 0.2, 0.4, 0.5, 0.4, 0.3, 0.5, 0.2\}$							

- [3] E. Fischetti ; A. Gisolfi; “From computer-aided instruction to intelligent tutoring systems”, Educ. Technol, vol, 30, no. 8, pp. 7–17, 1990.
- [4] 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, 1990.
- [5] J. Self, “The defining characteristics of intelligent tutoring systems research: ITSs care, precise”, Int. J. Artif. Intell. Educ, vol. 10, pp. 350–364, 1999.
- [6] R.Lelouche; “A collection of pedagogical agents for intelligent educational systems”, In intelligent tutoring systems:6<sup>th</sup> international Conference, ITS 2000: Lecture Notes in computer Science, pages 143–152. Springer-verlag, Berlin, 2000.
- [7] R. S. Legaspi ; R. C. Sison; “Modeling the tutor using reinforcement learning”, in Proc. PCSC, pp. 194–196, 2000.
- [8] P. Baffes ; R. Mooney; “Refinement-based student modeling and automated bug library construction”, J. Artif. Intell. Educ, vol. 7, no. 1, pp. 75–116, 1996.
- [9] 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., 1083-4419/\$26.00 © 2009 IEEE.
- [10] M. K. Hashem; “Learning automata based intelligent tutorial-like systems”, Ph.D. dissertation, School Comput. Sci., Carleton Univ., Ottawa, ON, Canada, 2007.
- [11] R.S. Sutton; A.C.Barto; “Reinforcement Learning:An introduvtion”, MIT Press, Cambridge, MA, Bradford Book, 1998A.
- [12] K. S. Narendra ; M. A. L. Thathachar; Learning Automata: An Introduction. Englewood Cliffs, NJ: Prentice-Hall, 1989.
- [13] M. A. L. Thathachar ; P. S. Sastry; “Varieties of learning automata: An overview”, IEEE Trans. Syst., Man, Cybern. B, Cybern., vol. 32, no. 6, pp. 711–722, Dec. 2002.

## ۷- جمع بندی

این مقاله استراتژی جدیدی برای سیستم شبیه آموزشی بیان می کند که به مدلسازی دانشجو اشاره دارد. مدل کننده دانشجو خودش از الگوریتم یادگیری Q به عنوان یک مکانیزم درونی برای تعیین مدل یادگیری دانشجو استفاده می کند بطوریکه می تواند در سیستم شبیه آموزشی جهت تعیین تمرين یادگیری برای هر دانشجو مورد استفاده قرار گیرد. برای رسیدن به این، مدل کننده دانشجو از الگوریتم یادگیری Q سطح بالا به نام Meta Q/LA می 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 یادگیری Meta Q/LA مشخص می کند که دانشجو در سوال، یادگیرنده کند، سریع و یا نرمال است.

این مقاله با استفاده از بکارگیری الگوریتم یادگیری Q سرعت همگرایی و در نتیجه تعیین مدل یادگیری دانشجو را بهبود داد و به سیستم شبیه آموزشی برای تعیین بهتر مدل دانشجو می تواند کمک 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 یادگیری Q در درصد بالایی از حالت ها، با دقت بالای ۹۰٪ مدل یادگیری را مشخص کرده است و تعداد تکرارها بطور چشمگیری کاهش داده شد.

## مراجع

- [1] V.J.Shute; J. Psotka; “Intelligent Tutoring Systems: Past, Present, and Future”, Handbook of Research on Educational Communications and Technology, Scholastic Publications, 1995.
- [2] E. Wenger; “Artificial Intelligence and Tutoring Systems: Computational and Cognitive Approaches to the Communication of Knowledge” ,Los Altos, CA: Morgan Kaufmann Publishers, Inc, 1987.

- [14] S. Lakshmivarahan; “*Learning Algorithms Theory and Applications*” New York: Springer-Verlag, 1981.
- [15] 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 2007. LNCS (LNAI), vol. 4570, pp. 813–822, Springer, Heidelberg (2007).
- [16] M.K., Hashem; B.J. Oommen; “*Learning Automata Based Intelligent Tutorial-like System*”, J.D. Velásquez et al. (Eds.): KES 2009, Part I, LNAI 5711, pp. 361–374, 2009. © Springer-Verlag Berlin Heidelberg 2009.
- [17] M.K., Hashem; B.J. Oommen; “*Using learning automata to model a domain in a tutorial-like system*”, In: Proceedings of ICMLC 2007, the 2007 International Conference of Machine Learning and Cybernetics, Hong Kong, pp.112–118, August 2007.

---

<sup>1</sup>Pedagogical

<sup>2</sup>Discrete Linear Reward Inaction

<sup>3</sup>Action

<sup>4</sup>fixed state transition probabilities

<sup>5</sup>Variable Structure Learning Automata

<sup>6</sup>Tutorial-like system

<sup>7</sup>Socratics

<sup>8</sup>Exploration

<sup>9</sup>Exploitive