

---

## بهبود عملکرد عامل‌ها در جامعه مصنوعی با استفاده از یادگیری تقویتی

---

امیرپویان خدابخشی<sup>۱</sup>، دکتر آرش رحمان<sup>۲</sup>، دکتر محسن روحانی<sup>۳</sup>

### چکیده

معمولاً در سیستم‌های چندعاملی، تعاملات بین عامل‌ها و تعاملات عامل‌ها با محیط، به صورت انتخاب و اجرای عمل‌هایی از بین مجموعه‌ای محدود از اعمال مشخص توسط عامل‌ها حاصل می‌شود. بنابراین نوع و میزان پیچیدگی رفتارهای پدیداری حاصل از این تعاملات نیز به نحوه اجرا و تعداد رفتارهای قابل اجرا توسط عامل‌ها وابسته است. در این پژوهش سعی شد با توسعه مدل کسب و انتقال تجربه و اضافه کردن قابلیت یادگیری به عامل‌ها، تاثیر یادگیری در بهبود رفتار عامل‌ها در انتخاب روش (راهبرد) های انتقال تجربه و در بهبود شاخص‌های رفاهی در جامعه مصنوعی مورد مطالعه قرار گیرد. روش یادگیری پیشنهاد شده در پژوهش برای افزایش دامنه توانایی‌های عامل‌ها، یادگیری تقویتی<sup>۴</sup> بود. با استفاده از این روش، عامل‌ها به مرور زمان یاد گرفتند که چگونه در مواجهه با شرایط مختلف محیطی، رفتارهای مناسب‌تری را انتخاب و اجرا کنند تا به اهداف فردی و اجتماعی نزدیک‌تر شوند. نتایج حاصل از شبیه‌سازی و انجام آزمایش‌ها نشان داد که اعمال فرآیند یادگیری می‌تواند منجر به بهبود رفتار عامل‌ها و بهبود شاخص‌های رفاهی جامعه مصنوعی شود.

**کلید واژه‌ها:** شبیه‌سازی اجتماعی، جامعه مصنوعی، مدل کسب و انتقال تجربه، یادگیری تقویتی

- 
۱. گروه مهندسی کامپیوتر، دانشکده فنی و مهندسی، واحد تهران جنوب، دانشگاه آزاد اسلامی، تهران، ایران.  
[amirpooyan@gmail.com](mailto:amirpooyan@gmail.com)
  ۲. گروه مهندسی کامپیوتر، دانشکده فنی و مهندسی، واحد تهران جنوب، دانشگاه آزاد اسلامی تهران، ایران.  
[arashrahman@yahoo.com](mailto:arashrahman@yahoo.com) نویسنده مسئول
  ۳. گروه مهندسی کامپیوتر، دانشکده فنی و مهندسی، واحد تهران جنوب، دانشگاه آزاد اسلامی، تهران، ایران.  
[m.rohani@niopdc.ir](mailto:m.rohani@niopdc.ir)

#### 4. Reinforcement Learning

## مقدمه

توسعه جوامع بشری، دانشمندان و محققان حوزه علوم اجتماعی را با چالش‌ها و مسائل پیچیده‌ای مواجه کرده است. مواجهه با پیچیدگی‌های جوامع بشری و لزوم ارائه راهکارهای مؤثر و بعضاً غیرقطعی، ضرورت وجود روش‌ها و ابزارهای نوین را برای آزمایش ایده‌ها، نظریه‌ها و راهکارهایی که دانشمندان به ازای چالش‌های مختلف اجتماعی مطرح می‌کنند، به وجود می‌آورد. شبیه‌سازی اجتماعی به عنوان روشی مؤثر امکان آزمایش و بررسی ایده‌ها و راه کارهایی که بعضاً مطالعه آنها در دنیای واقعی غیرممکن بوده یا مستلزم صرف هزینه‌های بسیاری است را فراهم آورده است. در این روش با ایجاد جوامع مصنوعی شامل محیط و (افراد جامعه در قالب) عامل‌های هوشمند، و تعیین و تنظیم پارامترها و قوانین محیطی و عاملی، سعی بر مطالعه پدیده‌ها و مسائل پیچیده اجتماعی است (خدابخشی و رحمان، ۱۳۹۲؛ رحمان و ستایشی، ۱۳۸۶).

در این مقاله این سؤال مطرح است که آیا استفاده از روش یادگیری توسط عامل‌ها در یک جامعه مصنوعی می‌تواند نقش مؤثری در بهبود رفتار آن‌ها (در انتخاب روش‌های انتقال تجربه) و به دنبال آن بهبود وضعیت رفاهی جامعه داشته باشد؟ برای پاسخ دادن به این سؤال سعی شد تا با در نظر گرفتن جامعه‌ای مصنوعی، تأثیر اجرای روش یادگیری تقویتی (سوتن<sup>۱</sup> و بارتو<sup>۲</sup>، ۱۹۹۸) توسط عامل‌ها در بهبود عملکرد آنها بررسی شود. به همین منظور، با اضافه کردن قابلیت یادگیری عامل در مدل کسب و انتقال تجربه (خدابخشی و رحمان، ۱۳۹۲) که بر مبنای مدل توزیع ثروت (لی<sup>۳</sup> و ویلسنکی<sup>۴</sup>، ۲۰۰۹) و مبتنی بر مدل فضای قندی طراحی شده و شامل شاخص‌های رفاهی است، تأثیر به‌کارگیری مکانیزم یادگیری مورد مطالعه قرار گرفت.

ایده این است که در فرایند یادگیری تقویتی، عامل‌ها به ازای اجرای عمل انتخاب کرده در وضعیت‌های مختلف، می‌توانند پاداش مثبت یا منفی دریافت کنند. بدین ترتیب با اجرای عمل‌های بیشتر توسط عامل‌ها و ثبت میزان پاداش متناظر با هر عمل در وضعیت‌های مختلف، به مرور زمان عامل‌ها یاد خواهند گرفت که در طول زندگی خود مناسب‌ترین عمل را در مواجهه با هر وضعیت محیطی انتخاب و اجرا کنند. در نتیجه با هدف بیشینه کردن میزان پاداش‌های دریافتی، قابل انتظار است که عامل‌ها بتوانند رفتار خود و به دنبال آن وضعیت رفاهی جامعه را

1. Sutton

2. Barto

3. Li

4. Wilensky

بهبود بخشند.

## مفاهیم پایه

### عامل<sup>۱</sup>

از دیدگاه هوش مصنوعی، عامل مؤلفه‌ای، پردازشی است که به طور هوشمندانه‌ای با توجه به شرایط محیطی که در آن قرار دارد، به صورت نسبتاً مستقل تصمیم‌گیری کرده و اعمال مختلفی را در جهت رسیدن به اهداف سیستم اجرا می‌کند. یک عامل نیاز به حس گرهایی دارد که به وسیله آن‌ها وضعیت خود و محیط را درک کند و براساس آن از میان عمل‌های ممکن، یک یا چند عمل را به عنوان اعمال مناسب برای مواجهه با شرایط جاری اجرا کند (وودریج<sup>۲</sup> و جینگز<sup>۳</sup>، ۱۹۹۵؛ سلرز<sup>۴</sup>، ۲۰۰۵؛ تیتن<sup>۵</sup>، ۲۰۱۰).

### یادگیری تقویتی

یادگیری تقویتی یک روش یادگیری است که در آن فرآیند یادگیری بر اساس بازخورد نتایج اقداماتی که عامل یادگیرنده در طول فعالیت خود انجام می‌دهد، هدایت می‌شود (کالبینگ و همکاران<sup>۶</sup>، ۱۹۹۶). در هر مرحله از فعالیت عامل، بازخورد اقدامات عامل در مواجهه با موقعیت‌های مختلف محیطی به صورت سیگنال‌های تشویقی (بازخورد مثبت) یا تنبیهی (بازخورد منفی) به عامل منتقل می‌شود و عامل سعی می‌کند با بهبود رفتار خود، مقدار پاداش دریافتی را بیشینه کند (سوتن و بارتو، ۱۹۹۸).

### شبیه‌سازی اجتماعی

با افزایش پیچیدگی ساختارها و مسائل جوامع بشری، نیاز به استفاده از روش‌های مؤثرتر و کم هزینه‌تر در انجام تحقیقات مرتبط با علوم اجتماعی اجتناب ناپذیر شده است. هر چند شیوه‌های استقرایی<sup>۷</sup> و قیاسی<sup>۸</sup> روش‌های مؤثر و متداول تحقیق و مطالعه پدیده‌های اجتماعی هستند، ولیکن برخی از دانشمندان معتقدند که شبیه‌سازی می‌تواند به عنوان روش سوم در تحقیقات این شاخه از علم مورد استفاده قرار گیرد (آکسلرود<sup>۹</sup>، ۱۹۹۷).

1. Agent

2. Wooldridge, M.

3. Jennings, N. R.

4. Sellers, B.

5. Teaten, W. J.

6. Kaelbling et al.

7. Inductive

8. Deductive

9. Axelrod

در روش شبیه‌سازی اجتماعی با استفاده از ابزارها و روش‌های رایانه‌ای، پدیده‌های اجتماعی شبیه‌سازی می‌شوند. به این ترتیب، شبیه‌سازی کامپیوتری می‌تواند برای بررسی و تحلیل پدیده‌های اجتماعی، پیش‌بینی رفتارهای اجتماعی و آزمایش برخی از نظریه‌ها مورد استفاده قرار گیرد.

#### مدل فضای قندی

مدل فضای قندی، مدلی برای شبیه‌سازی و مطالعه فرآیندهای اجتماعی و رفتارهای پیدایشی ناشی از تعاملات اجتماعی است (اپستین<sup>۱</sup> و آکستل<sup>۲</sup>، ۱۹۹۶). این مدل، متشکل از عامل‌ها، محیط و مجموعه‌ای از قوانین است که این قوانین، نحوه رفتارهای فردی عامل‌ها و تعاملات عامل‌ها با یکدیگر را تعیین می‌کنند. همچنین نحوه ارتباط و تأثیر فعالیت عامل‌ها بر محیط نیز توسط مجموعه قوانین بیان می‌شود (اپستین و آکستل، ۱۹۹۶).

محیط در مدل فضای قندی، از شبکه‌ای از سلول‌ها تشکیل شده است. هر سلول می‌تواند دارای مقداری قند باشد. عامل‌ها برای ادامه بقای خود به قند احتیاج دارند و بنابراین هدف عامل‌ها در مدل فضای قندی، برداشت قند موجود در سلول‌های محیط است. قوانین محیطی تعیین می‌کنند که عامل‌ها تحت چه شرایطی می‌توانند قند موجود در سلول‌ها را برداشت کنند. به این ترتیب عامل‌ها می‌توانند در محیط حرکت کرده و طبق قوانین از منابع محیطی برای ادامه بقا خود استفاده کنند (اپستین و آکستل، ۱۹۹۶).

از مدل فضای قندی برای توصیف و مطالعه بسیاری از پدیده‌های اجتماعی استفاده شده است (اپستین و آکستل، ۱۹۹۶). در این مدل، عامل‌ها در هر بازه زمانی از میان تعدادی از عمل‌های از پیش تعیین شده، یک یا چند عمل را با توجه به قوانین حاکم بر جامعه شبیه‌سازی شده انتخاب کرده و اجرا می‌کنند. به این طریق می‌توان، با اندازه‌گیری شاخص‌های مختلف وابسته به عامل‌ها و محیط، تأثیر بروز رفتارهای مختلف از سوی عامل‌ها را مطالعه کرد.

#### مدل کسب و انتقال تجربه

مدل کسب و انتقال تجربه، با اضافه کردن قابلیت کسب و انتقال تجربه به مدل توزیع ثروت (لی و ویلنسکی، ۲۰۰۹) ایجاد شده است (خدابخشی و رحمان، ۱۳۹۲).

مدل توزیع ثروت (لی و ویلنسکی، ۲۰۰۹)، مدلی مبتنی بر مدل فضای قندی بوده و در آن

1. Epstein
2. Axtell

عامل‌ها با پیروی از قوانین از پیش تعیین شده، اقدام به جمع‌آوری قند از منابع محیطی می‌کنند (لی و ویلنسکی، ۲۰۰۹). در مدل توزیع ثروت، شاخص ضریب جینی و منحنی لورنز به منظور نمایش نحوه توزیع درآمد یا ثروت مورد استفاده قرار گرفته است.

قابلیت کسب تجربه به عامل‌ها کمک می‌کند که اطلاعاتی را راجع به محیطی که در آن فعالیت می‌کنند ثبت کرده و در ادامه روند فعالیت‌های خود در محیط، از آن اطلاعات استفاده کنند. بدین ترتیب مکانیزمی برای کاهش فعالیت تصادفی عامل‌ها ایجاد می‌شود و عامل‌ها با هوشمندی بیشتری در محیط فعالیت می‌کنند (خدابخشی و رحمان، ۱۳۹۲).

ویژگی‌های منحصر به فرد محیط در مدل کسب و انتقال تجربه نسبت به مدل توزیع ثروت، شامل توزیع تصادفی قند در محیط و وجود یک کلید منحصر به فرد برای هر سلول است.

در مدل کسب و انتقال تجربه، عامل‌ها برای اینکه بتوانند قند موجود در سلول‌ها را برداشت کنند باید کلید متناظر با سلول مورد نظر را در اختیار داشته باشند که این امر مستلزم مجهز شدن عامل‌ها به حافظه‌ای برای ثبت کلیدهای کشف شده است. وقتی یک عامل می‌خواهد قند موجود در یک سلول را برداشت می‌کند، ابتدا به حافظه خود مراجعه کرده و مختصات سلول مورد نظر و کلید متناظر با آن را در حافظه خود جستجو می‌کند. در صورتی که کلید یافت شود، آن را به سلول ارائه کرده و محتویات سلول را برداشت کند. لیکن در صورتی که کلید متناظر با سلول را در حافظه خود نداشته باشد، یک کلید تصادفی تولید می‌کند و آن را به سلول مورد نظر ارائه می‌کند. در صورتی که کلید تولید شده، کلید سلول باشد، ضمن برداشت قند موجود در سلول، آن کلید را به عنوان یک تجربه در حافظه خود ثبت کرده و در مراجعات بعدی به همان سلول، از کلید ذخیره شده استفاده می‌کند (خدابخشی و رحمان، ۱۳۹۲).

در شرایط مناسب، عامل‌ها کلیدهایی را که در اثر فعالیت در محیط کشف کرده‌اند، با استفاده از راهبردهای انتقال تجربه (خدابخشی و رحمان، ۱۳۹۲) می‌توانند به سایر عامل‌ها منتقل کنند.

فرآیند انتقال در راهبرد انتقال تجربه مشارکتی<sup>۱</sup> به این صورت است که در یک گروه محلی، ابتدا تمامی عامل‌ها، تجربیات خود را به صورت یک مجموعه واحد روی هم گذاشته و تجمیع می‌کنند. سپس همان عامل‌ها، مجموعه تجمیع شده تجربیات را به حافظه خود منتقل کرده و مورد بهره‌برداری قرار می‌دهند.

در راهبرد انتقال تجربه از تمامی عامل‌ها به کم تجربه‌ترین عامل(ها)، تجربیات همه اعضا گروه(های) محلی جمع می‌شود. لیکن فقط آن عامل(هایی) که نسبت به بقیه تعداد تجربیات کمتری در حافظه خود دارد(دارند)، تجربیات جمع شده را به حافظه خود منتقل می‌کند (می‌کنند).

در راهبرد انتقال تجربه از با تجربه‌ترین عامل(ها) به تمامی عامل‌ها، آن عامل(هایی) که نسبت به بقیه اعضای گروه(های) محلی دارای تعداد بیشتری تجربه در حافظه تجربیات خود است(هستند) یعنی عامل(های) باتجربه‌تر، تجربیات خود را به بقیه عامل‌های همسایه انتقال می‌دهد (می‌دهند).

در راهبرد انتقال تجربه از با تجربه‌ترین عامل(ها) به فقیرترین عامل(ها)، با تجربه‌ترین عامل(ها) تجربیات خود را به عامل(های) فقیر موجود در همسایگی خود انتقال می‌دهد (می‌دهند). اعضا فقیر، اعضای هستند که میزان قند اندوخته (ثروت) آنها از دیگر اعضا کمتر است.

در راهبرد انتقال تجربه از تمامی عامل‌ها به فقیرترین عامل(ها)، تجربیات همه اعضا جمع شده و فقیرترین عامل(ها)، مجموعه تجربیات جمع شده را به حافظه خود منتقل کرده و مورد استفاده قرار می‌دهد (می‌دهند).

اجرای راهبردهای فوق الذکر و بررسی تأثیر آن‌ها بر وضعیت زندگی عامل‌ها در محیط، مشخص کرد که برخی از راهبردها، تأثیر بهتری در بهبود وضعیت زندگی و بقاء عامل‌ها دارند (خدابخشی و رحمان، ۱۳۹۲).

در مدل کسب و انتقال تجربه، عامل‌ها موظف به اجرای راهبردهای از پیش تعیین شده‌اند و در انتخاب راهبرد انتقال تجربه مختار نیستند.

### مدل سازی جامعه یادگیرنده

در پژوهش، توسعه مدل کسب و انتقال تجربه مدنظر قرار گرفت و امکان یادگیری به مجموعه قابلیت‌های عامل‌ها اضافه شد و مدلی تحت عنوان مدل جامعه یادگیرنده شکل گرفت. روش یادگیری مورد استفاده در این مدل روش یادگیری تقویتی بود. بدین ترتیب، در این مدل، انتخاب راهبردهای مختلف انتقال تجربه در شرایط مختلف زندگی عامل‌ها به اختیار خود آنها صورت می‌گرفت.

ویژگی‌های مدل جامعه یادگیرنده

محیط در مدل جامعه یادگیرنده، دارای خواص و ویژگی‌های مشابه با محیط در مدل کسب و انتقال تجربه است. عامل‌ها در مدل جامعه یادگیرنده دارای ویژگی‌های زیر هستند:

الف. بر خلاف مدل کسب و انتقال تجربه که در آن عامل‌ها ملزم به اجرای یک راهبرد انتقال تجربه از پیش تعیین شده هستند، عامل‌ها در مدل جامعه یادگیرنده، در انتخاب راهبرد مورد نظر خود از بین پنج راهبرد انتقال تجربه پیشنهاد شده مختار هستند.

ب. وضعیت هر عامل نسبت به سایر عامل‌ها براساس دو پارامتر ثروت و تجربه سنجیده می‌شود. به بیان دیگر، در هر مرحله از اجرای فرآیند یادگیری، وضعیت اجتماعی هر عامل جهت تعیین میزان بهبودی که فرآیند انتقال تجربه در زندگی عامل ایجاد می‌کند، مورد ارزیابی قرار می‌گیرد. به همین منظور، عامل‌ها گروه‌های محلی تشکیل داده و وضعیت هر یک از اعضا گروه، نسبت به ثروتمندترین و باتجربه‌ترین اعضاء گروه سنجیده می‌شود.

ج. در مدل جامعه یادگیرنده، عامل‌ها علاوه بر حافظه‌ای برای ذخیره تجربیات خود، حافظه‌ای برای ذخیره ارزش هر یک از راهبردهای انتقال تجربه به ازای وضعیت‌های مختلف اجتماعی دارند. حافظه راهبردها، مجموعه‌ای است از سه تایی‌های مرتب به شکل  $\langle S, P, R \rangle$  که در آن  $S$  وضعیت ارزش‌گذاری شده،  $P$  عمل انجام شده به ازای وضعیت  $S$  و  $R$  ارزش محاسبه شده برای  $P$  در مواجهه با  $S$  می‌باشد.

ارزش‌گذاری راهبردها

اجرای هر راهبرد انتقال تجربه، تغییراتی را در روند زندگی عامل ایجاد می‌کند که با محاسبه میزان بهبودی یا تغییری که اجرای راهبرد انتقال تجربه در وضعیت جاری زندگی عامل ایجاد کرده است، ارزش تأثیر راهبرد انتقال تجربه به دست می‌آید.

از آنجا که هدف عامل‌ها در هر دو مدل کسب و انتقال تجربه و جامعه یادگیرنده، جمع‌آوری قند بیشتر از محیط است، ارزش راهبردهای انتقال تجربه براساس تأثیری که بر بهبود توانایی عامل‌ها در جمع‌آوری قند از محیط می‌گذارند، سنجیده می‌شود.

با ارزش‌گذاری راهبردهای انتقال تجربه در مواجهه با وضعیت‌های مختلف، وقتی که عامل مجدداً در همان وضعیت قرار بگیرد، قادر به انتخاب مؤثرترین راهبرد از میان راهبردهای انتقال تجربه است. عامل‌ها، به طور مداوم مقادیر ارزش راهبردها به ازای وضعیت‌های مختلف را به روز رسانی می‌کنند و به مرور زمان، مقادیر دقیق‌تری را به عنوان ارزش راهبردها به دست می‌آورند.

استفاده از روش  $\epsilon$ -greedy در ارزش‌گذاری راهبردها

اگر یک عامل برای اولین بار با وضعیتی مواجه شود، به صورت تصادفی یکی از راهبردهای انتقال تجربه را انتخاب و اجرا کرده و پس از اجرای آن، مقدار ارزش‌گذاری شده متناظر با وضعیت و راهبرد اجرا شده را در حافظه راهبردهای خود ذخیره می‌کند. در صورتی که عامل مجدداً با همان وضعیت مواجه شود، به حافظه راهبردهای خود مراجعه کرده و با توجه به اینکه راهبردی را متناظر با وضعیت جاری پیدا می‌کند، آن را اجرا می‌کند. در این صورت، ممکن است راهبردهای دیگری وجود داشته باشند که در مواجهه با وضعیت جاری کارایی بهتری داشته باشند و عامل، آنها را اجرا نکند. به منظور جلوگیری از بروز چنین حالتی، عامل‌ها در مواجهه با وضعیت‌های مختلف، با احتمال  $\epsilon$  که مبین احتمال کوچکی است، از بین راهبردهای دیگر که تاکنون در مواجهه با وضعیت جاری اجرا نکرده‌اند، یکی را انتخاب و اجرا می‌کنند. به این ترتیب، راهبردهایی که در مواجهه با وضعیت‌های مختلف دارای تأثیر بهتری هستند هم کشف و اجرا می‌شوند.

مشابهت وضعیت‌ها

با توجه به اینکه عامل‌ها به طور پیوسته در محیط حرکت می‌کنند، ممکن است در گروه‌های محلی مختلفی قرار بگیرند و متعاقب آن در وضعیت‌هایی قرار بگیرند که قبلاً در وضعیت‌هایی مشابه آن واقع شده باشند. در چنین شرایطی به نظر می‌رسد که انتخاب راهبردهایی که قبلاً در مواجهه با وضعیت‌های مشابه اجرا شده و ارزش‌گذاری شده‌اند، نسبت به انتخاب تصادفی راهبردها تأثیر بهتری داشته باشند. به همین دلیل، عامل‌ها در صورتی که با شرایطی مشابه با آنچه که قبلاً تجربه کرده‌اند مواجه شوند، راهبردی را که در مورد وضعیت مشابه بالاترین کارایی را داشته است، انتخاب و اجرا می‌کنند.

تعیین مشابهت وضعیت‌ها در مدل جامعه یادگیرنده با استفاده از تعیین فاصله اقلیدسی بردار وضعیت جاری و بردار وضعیت قبلی انجام می‌شود. در صورتی که فاصله اقلیدسی بردارهای دو وضعیت از مقداری مانند  $d$  کمتر باشد، آن دو وضعیت توسط عامل مشابه تشخیص داده می‌شوند و راهبرد متناظر با آن اجرا می‌شود. پس از اجرای راهبرد متناظر با وضعیت مشابه و محاسبه ارزش اجرای آن در مواجهه با وضعیت جاری، مدخل هر دو وضعیت در حافظه راهبردها به روز رسانی می‌شود.



شاخص‌های رفاهی

در مدل کسب و انتقال تجربه، به منظور مطالعه رفتار مدل و عملکرد عامل‌ها از شاخص‌های ضریب جینی، میانگین ثروت عامل‌ها و میزان مرگ و میر ناشی از گرسنگی استفاده شد (خدابخشی و رحمان، ۱۳۹۲). در مدل جامعه یادگیرنده هم از همان شاخص‌ها استفاده شده است.

### آزمایش‌ها و تحلیل یافته‌ها

با استفاده از ابزار نرم‌افزاری NetLogo (ویلنسکی، ۱۹۹۹)، جامعه مصنوعی یادگیرنده شبیه سازی و آزمایش‌هایی با آن انجام شد. سعی شد تا با مقایسه یافته‌های حاصل از آزمایش‌ها، عملکرد عامل‌ها با وجود یادگیری تقویتی بررسی شود.

آزمایش‌های اول و دوم

آزمایش اول با مدل جامعه مصنوعی یادگیرنده در شرایط غیرفعال بودن قابلیت یادگیری برای عامل‌ها صورت گرفت و عامل‌ها در هر مرحله از زندگی خود به منظور انتقال تجربیات خود به یکدیگر، به طور تصادفی (با توزیع یکنواخت) از بین راهبردهای انتقال تجربه یکی را انتخاب و اجرا می‌کردند.

آزمایش دوم با جامعه مصنوعی یادگیرنده در شرایط فعال بودن قابلیت یادگیری برای عامل‌ها صورت گرفت و در آن عامل‌ها با استفاده از روش یادگیری تقویتی، راهبردهای انتقال تجربه را ارزش‌گذاری، انتخاب و اجرا می‌کردند.

در هر دو آزمایش فوق، تعداد دفعاتی که هر یک از راهبردهای انتقال تجربه توسط عامل‌ها انتخاب و اجرا شدند، در قالب نمودارهایی ترسیم شد. سپس یافته‌های حاصل از آزمایش دوم با یافته‌های حاصل از آزمایش اول مقایسه شد.

در انجام آزمایش‌ها، مقادیر پارامترهای محیطی و عاملی در نظر گرفته شده در طول تمامی آزمایش‌ها یکسان بود. این مقادیر در جداول ۱ و ۲ مشخص شده است.

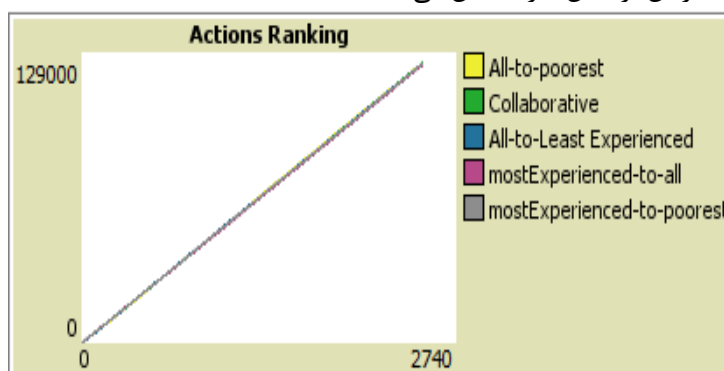
### جدول ۱: مقادیر مربوط به پارامترهای محیطی در آزمایش‌های اول و دوم

میزان پیچیدگی محیط	حداکثر ظرفیت قندی هر سلول	کلید سلول‌ها
۳۵٪	۵۰ واحد قند	تصادفی مابین ۱ تا ۳۵

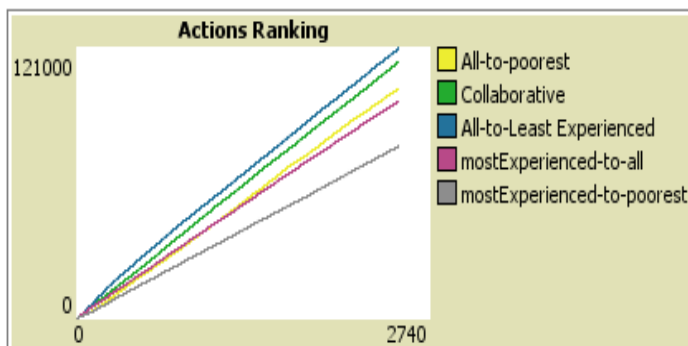
جدول ۲: مقادیر مربوط به پارامترهای عاملی در آزمایش‌های اول و دوم

وسعت دید	حداکثر طول عمر عامل‌ها	میزان متابولیسم عامل‌ها	حداکثر ثروت اولیه هر عامل	حداقل ثروت اولیه هر عامل	جمعیت اولیه عامل‌ها
تصادفی بین ۶ تا ۱	تصادفی بین ۶۰ تا ۱۰۰	تصادفی بین ۱ تا ۴	۲۵ واحد قند	۵ واحد قند	۲۵۰ عامل

در آزمایش‌ها، ابتدا شمارنده‌هایی متناظر با هر یک از راهبردهای انتقال تجربه تعریف شد. با شروع آزمایش، مقدار هر یک از شمارنده‌ها برابر با صفر بود و هر بار که یکی از راهبردهای انتقال تجربه توسط یکی از عامل‌ها انتخاب می‌شد، یک واحد به شمارنده متناظر با آن راهبرد اضافه می‌شد. بدین ترتیب با ثبت لحظه به لحظه مقادیر شمارنده‌ها در طول آزمایش‌ها، نمودارهایی رسم شد که نشان دهنده میزان استقبال عامل‌ها از هر یک از راهبردهای انتقال تجربه و یا به عبارتی رتبه‌بندی راهبردهای انتقال تجربه در انتخاب توسط عامل‌ها است. نتایج آزمایش‌های اول و دوم: بهبود رفتار عامل‌ها در انتخاب روش (راهبرد)‌های انتقال تجربه با وجود قابلیت یادگیری اشکال زیر نمودارهای مربوط به تعداد دفعات انتخاب راهبردهای انتقال تجربه توسط عامل‌ها در طول مدت زمان آزمایش‌ها را نمایش می‌دهند.



شکل ۱: نمودارهای مربوط به تعداد دفعات انتخاب راهبردهای انتقال تجربه به صورت تصادفی توسط عامل‌ها در آزمایش اول (در حالت عدم وجود قابلیت یادگیری عامل)



شکل ۲: نمودارهای مربوط به تعداد دفعات انتخاب راهبردهای انتقال تجربه توسط عامل‌ها در آزمایش دوم (با وجود قابلیت یادگیری تقویتی)

در شکل ۲ نمودار آبی رنگ مربوط به راهبرد انتقال تجربه تمامی عامل‌ها به کم تجربه ترین عامل(ها)، نمودار سبز رنگ مربوط به راهبرد انتقال تجربه مشارکتی، نمودار زرد رنگ مربوط به راهبرد انتقال تجربه تمامی عامل‌ها به فقیرترین عامل(ها)، نمودار بنفش رنگ مربوط به راهبرد انتقال تجربه باتجربه ترین عامل(ها) به تمامی عامل‌ها و نمودار خاکستری رنگ مربوط به راهبرد انتقال تجربه با تجربه ترین عامل(ها) به فقیرترین عامل(ها) است.

شکل ۱ نشان می‌دهد که وقتی عامل‌ها بدون استفاده از یادگیری اقدام به انتخاب راهبردهای انتقال تجربه می‌کنند، تمایزی بین راهبردها و نتایج به دست آمده از آنها قائل نمی‌شوند. لیکن در شکل ۲ مشاهده می‌شود که وقتی عامل‌ها از یادگیری برای انتخاب راهبردهای انتقال تجربه استفاده می‌کنند، در تعداد دفعات انتخاب و اجرای راهبردها تفاوت‌هایی ایجاد می‌شود. به بیان دیگر، برخی از راهبردهای انتقال تجربه بیشتر و برخی دیگر کمتر، توسط عامل‌ها انتخاب و اجرا می‌شوند.

در شکل ۲ مشاهده می‌شود که راهبرد انتقال تجربه تمامی عامل‌ها به کم تجربه ترین عامل(ها) (نمودار آبی رنگ) نسبت به سایر راهبردهای انتقال تجربه، بیشتر توسط عامل‌ها انتخاب و اجرا شده است. در واقع این راهبرد توسط عامل‌ها بیشتر در وضعیت‌های مختلف مناسب‌تر تشخیص داده شده و لذا بیشتر انتخاب و اجرا شده است. به همین ترتیب نمودارها نشان می‌دهند که راهبردهای انتقال تجربه مشارکتی، انتقال تجربه همه عامل‌ها به فقیرترین عامل(ها)، انتقال تجربه با تجربه ترین عامل(ها) به تمامی عامل‌ها و انتقال تجربه با تجربه ترین

عامل(ها) به فقیرترین عامل(ها) از نظر دفعات انتخاب و اجرا توسط عامل‌ها در رتبه‌های دوم تا پنجم قرار دارند. رتبه‌بندی انتخاب راهبردهای انتقال تجربه توسط عامل‌ها با وجود قابلیت یادگیری با توجه به نتایج بدست آمده از آزمایش دوم در جدول ۳ مشخص شده است.

جدول ۳: رتبه‌بندی راهبردهای انتقال تجربه به لحاظ تعداد دفعات انتخاب و اجرا توسط عامل‌ها

رتبه از نظر انتخاب و اجرا توسط عامل‌ها	رنگ نمودار	راهبرد انتقال تجربه
اول	آبی	تمامی عامل‌ها به کم تجربه‌ترین عامل(ها)
دوم	سبز	مشارکتی
سوم	زرد	تمامی عامل‌ها به فقیرترین عامل(ها)
چهارم	بنفش	باتجربه‌ترین عامل(ها) به تمامی عامل‌ها
پنجم	خاکستری	باتجربه‌ترین عامل(ها) به فقیرترین عامل(ها)

در جدول ۴ نتایجی از آزمایش‌های انجام شده با مدل کسب و انتقال تجربه در ارتباط با میزان بهبود شاخص‌های ضریب جینی و میانگین ثروت عامل‌ها در حالت اجرای به تفکیک راهبردهای انتقال تجربه مشخص شده است. در این جدول راهبردهای انتقال تجربه بر اساس میزان بهبود وضعیت رفاهی مرتب شده‌اند. همان‌طور که مشاهده می‌شود، راهبرد انتقال تجربه تمامی عامل‌ها به کم تجربه‌ترین عامل (ها) و بعد از آن مشارکتی به لحاظ دو شاخص میزان بهبودی ضریب جینی و میزان بهبودی میانگین ثروت بیشترین اثربخشی را در بهبود وضعیت رفاهی داشته‌اند.

جدول ۴: میزان تأثیر به کارگیری راهبردهای انتقال تجربه در بهبود شاخص‌های رفاهی

میزان بهبودی ضریب جینی	میزان بهبودی میانگین ثروت	راهبرد انتقال تجربه
۳۰/۵٪	۷۵۰٪	تمامی عامل‌ها به کم تجربه‌ترین عامل(ها)
۲۹/۹٪	۷۰۲٪	مشارکتی <sup>۱</sup>
۲۹/۹٪	۷۰۰٪	تمامی عامل‌ها به فقیرترین عامل(ها)
۹/۳٪	۳۴۶٪	با تجربه‌ترین عامل(ها) به تمامی عامل‌ها
۵/۵٪	۲۲۲٪	با تجربه‌ترین عامل(ها) به فقیرترین عامل(ها)

مقایسه جدول های ۳ و ۴ نشان می‌دهد که رتبه‌بندی نتیجه شده از آزمایش دوم این بخش، با رتبه بندی حاصل از نتایج آزمایش‌های انجام شده با مدل کسب و انتقال تجربه مطابقت دارد. نتایج نشان می‌دهد که راهبردهایی که در آزمایش‌های مربوط به مدل کسب و انتقال تجربه اثربخشی بیشتری در بهبود شاخص‌های رفاهی از خود نشان داده‌اند، در آزمایش دوم (در حالت وجود قابلیت یادگیری) نیز در دفعات بیشتر و رتبه بالاتری توسط عامل‌ها انتخاب و اجرا شده‌اند. به عبارت دیگر رفتار عامل‌ها در انتخاب روش‌های انتقال تجربه با وجود قابلیت یادگیری بهبود یافته است.

۱. یافته‌های به دست آمده از آزمایش‌های تکمیلی با مدل کسب و انتقال تجربه نشان داد که راهبرد مشارکتی و راهبرد انتقال تجربه تمامی عامل‌ها به فقیرترین عامل(ها) اثرگذاری نسبتاً مشابهی در بهبود میانگین ثروت عامل‌ها دارند. نتایج برخی از آزمایش‌های گذشته نیز اثرگذاری اندک بیشتر راهبرد انتقال تجربه تمامی عامل‌ها به فقیرترین عامل(ها) را نسبت به راهبرد مشارکتی در بهبود میانگین ثروت عامل‌ها نشان داده است (نتایج آزمایش در مقاله (خدابخشی و رحمان، ۱۳۹۲)). از آنجائی که راهبرد مشارکتی اثرگذاری بیشتری در بهبود میزان مرگ و میر ناشی از گرسنگی و حتی ضریب جینی نسبت به راهبرد انتقال تجربه تمامی عامل‌ها به فقیرترین عامل(ها) از خود نشان داده است، لذا ارجحیت آن در بهبود وضعیت رفاهی عامل‌ها نسبت به راهبرد انتقال تجربه تمامی عامل‌ها به فقیرترین عامل(ها) قابل ذکر است.

آزمایش‌های سوم و چهارم

آزمایش سوم با مدل جامعه مصنوعی یادگیرنده در شرایط غیرفعال بودن قابلیت یادگیری برای عامل‌ها صورت گرفت و عامل‌ها در هر مرحله از زندگی خود به منظور انتقال تجربیات خود به یکدیگر، به طور تصادفی از بین راهبردهای انتقال تجربه یکی را انتخاب و اجرا می‌کردند.

آزمایش چهارم با مدل جامعه مصنوعی یادگیرنده در شرایط فعال بودن قابلیت یادگیری برای عامل‌ها صورت گرفت و در آن عامل‌ها با استفاده از روش یادگیری تقویتی، راهبردهای انتقال تجربه را ارزش‌گذاری، انتخاب و اجرا می‌کردند.

آزمایش‌های فوق تا ۲۵۰۰ واحد زمانی انجام و به تعداد ۱۰ بار تکرار شدند. این بار در پایان هر آزمایش، شاخص‌های رفاهی جامعه مصنوعی اندازه‌گیری و ثبت شد و به منظور به دست آوردن نتایج دقیق‌تر از آن مقادیر میانگین گرفته شد. نهایتاً نتایج آزمایش‌ها با یکدیگر مقایسه شد و تاثیر استفاده از یادگیری در بهبود وضعیت شاخص‌های رفاهی جامعه مصنوعی مورد بررسی قرار گرفت.

نتایج آزمایش‌های سوم و چهارم: بهبود وضعیت شاخص‌های رفاهی جامعه مصنوعی با وجود قابلیت یادگیری

نتایج حاصل از آزمایش‌های سوم و چهارم در جدول ۵ نمایش داده شده است.

جدول ۵: میزان اثربخشی به کارگیری یادگیری در بهبود شاخص‌های رفاهی با وجود تمامی راهبردهای انتقال تجربه

میزان بهبود	مقدار میانگین در آزمایش‌ها		شاخص	آزمایش اول میانگین مقدار
	بدون به کارگیری یادگیری	با به کارگیری یادگیری		
٪ ۰/۳	۰/۳۷۴	۰/۳۷۳	ضریب جینی	
٪ ۰/۱۸	۲۴۱	۲۴۳	میانگین ثروت عامل‌ها	
٪ ۳/۳	۱۷۶۰	۱۷۰۳	مرگ و میر ناشی از گرسنگی	

مقادیر مندرج در جدول ۵ ظاهراً نشان می‌دهد که به کارگیری قابلیت یادگیری تأثیر چشم‌گیری در بهبود شاخص‌های رفاهی جامعه ندارد. دلایل متعددی برای کم اثر بودن به کارگیری فرآیند یادگیری در بهبود شاخص‌های رفاهی در آزمایش‌های این بخش قابل طرح است. مهم‌ترین دلیل مربوط به مساعد بودن شرایط محیطی برای تداوم زندگی عامل‌ها است. به بیان دیگر در طول اجرای آزمایش‌ها، شرایط زندگی عامل‌ها به اندازه‌ای دشوار نبوده که نقش یادگیری در زندگی آن‌ها پررنگ‌تر و برجسته‌تر شود و اهمیت استفاده از آن برای انتخاب راهبردهای مناسب‌تر ملموس‌تر شود.

با توجه به دلیل فوق سعی شد که شرایط محیط در هر دو آزمایش به شکلی تغییر کند که شرایط زندگی عامل‌ها نسبت به قبل دشوارتر شود و عامل‌ها برای مواجهه با شرایط محیطی با چالش‌ها و موانع بیشتری برای بقاء مواجه شوند. بدین ترتیب پیچیدگی محیط از ۳۵ درصد به ۶۰ درصد و متابولیسم عامل‌ها از مقادیر تصادفی بین ۱ تا ۴ به مقادیر تصادفی بین ۴ تا ۱۰ افزایش داده شد. در این شرایط فضای کلیدهای مربوط به سلول‌های محیط گسترده‌تر شده و شانس عامل‌ها برای یافتن کلیدهای متناظر با سلول‌ها به صورت تصادفی، کاهش پیدا می‌کند. بنابراین عامل‌ها مجبور به انتخاب و اجرای راهبردهای انتقال تجربه به شکل کارآمدتر هستند. همچنین افزایش متابولیسم عامل‌ها موجب می‌شود که عامل‌ها قندی که از فعالیت در محیط به دست آورده‌اند را با سرعت بیشتری مصرف کنند و لذا خطر گرسنگی و مرگ ناشی از آن در بین عامل‌ها افزایش پیدا می‌نماید. در چنین شرایطی در اختیار داشتن تجربیات لازم برای بهره‌برداری از قند موجود در سلول‌ها اهمیت و ارزش بیشتری پیدا کرده و بنابراین انتخاب اثربخش‌ترین راهبردهای انتقال تجربه در جهت بقاء عامل‌ها اهمیت بیشتری دارد.

بدین ترتیب با تغییر مقدار دو پارامتر پیچیدگی محیط و بازه متابولیسم عامل‌ها، مطابق با توضیحات فوق‌الذکر، مجدداً آزمایش‌های سوم و چهارم تکرار شدند. به منظور اطمینان بیشتر از یافته‌های حاصل از آزمایش‌ها، هر کدام از آزمایش‌ها ۲۰ بار تکرار شدند. یافته‌ها به صورت میانگین در جدول ۶ مشخص شده‌اند.

جدول ۶: میزان اثربخشی به کارگیری یادگیری در بهبود شاخص‌های رفاهی با وجود تشدید موانع برای بقای عامل‌ها

میزان بهبود	مقدار میانگین در آزمایش‌ها		شاخص
	بدون به کارگیری یادگیری	با به کارگیری یادگیری	
٪ ۰/۸	۰/۴۷۸	۰/۴۷۴	ضریب جینی
٪ ۲۸/۷	۳۸/۱۵	۵۳/۵	میانگین ثروت عامل‌ها
٪ ۲۳/۱	۱۳۷۱۴۳	۱۰۵۵۳۵	مرگ و میر ناشی از گرسنگی

بررسی یافته‌های مندرج در جدول ۶ نشان می‌دهد که در چنین شرایطی، با وجود یادگیری، بهبودی بیشتری در شاخص‌های میانگین ثروت عامل‌ها و مرگ و میر ناشی از گرسنگی حادث شده است.

### بحث و نتیجه‌گیری

نتایج آزمایش‌ها نشان داد که استفاده از قابلیت یادگیری تقویتی می‌تواند منجر به بهبود رفتار عامل‌ها در انتخاب روش (راهبرد)های انتقال تجربه شود که به دنبال آن بهبود شاخص‌های رفاهی جامعه مصنوعی را در پی خواهد داشت.

هم چنین مشخص شد که تشدید موانع بقای عامل‌ها در جامعه مصنوعی، می‌تواند ضرورت وجود یا لزوم استفاده از قابلیت یادگیری عامل را به منظور انتخاب راهبردهای مؤثرتر، محسوس‌تر کند. چرا که عامل‌ها در مواجهه با شرایط دشوار محیطی، به منظور ادامه بقا، مجبور به انجام رفتارهای هوشمندانه‌تر و کارآمدتر هستند که به نظر این مهم با وجود قابلیت یادگیری برای عامل‌ها قابل حصول است.



## منابع

### فارسی

- رحمان، آ. و ستایشی، س. (۱۳۸۶). نقش توزیع ثروت، وراثت و کنترل جمعیت در رفاه اجتماعی: شبیه سازی رفاه اجتماعی در جامعه مصنوعی. *فصلنامه رفاه اجتماعی*، ۷ (۲۶)، ۱۸۳-۲۰۰.
- خدابخشی، ا. و رحمان، آ. (۱۳۹۲). راهبردهای مختلف انتقال تجربه بین عامل ها و مقایسه تاثیرات اجرای آنها بر رفاه اجتماعی در جامعه مصنوعی. *فصلنامه پژوهش اجتماعی*، ۶ (۲۰)، ۱۳-۳۷.

### انگلیسی

- Axelrod, R. (1997). **Advancing the art of simulation in the social sciences**. *Complexity*, 3(2), pp. 16–22.
- Epstein, J. M., & Axtell, R. (1996). **Growing artificial societies: social science from the bottom up**. Brookings Institution Press.
- Kaelbling, L. P., Littman, M. L., & Moore, A. W. (1996). **Reinforcement learning: A survey**. *Journal of Artificial Intelligence Research*, 4 237–285.
- Li, J. & Wilensky, U. (2009). **NetLogo Sugarscape3 Wealth Distribution model**. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL. <http://ccl.northwestern.edu/netlogo/models/Sugarscape3WealthDistribution>
- Sellers, B. (2005). **Agent-oriented methodologies**. Hershey, PA: Idea Group Pub.
- Sutton, R. S., & Barto, A. G. (1998). **Reinforcement learning an introduction**. Cambridge, Mass.: MIT Press.
- Teaten, W. J. (2010). **Artificial Intelligence- Agents and Environments**. BookBoon.
- Wilensky, U. (1999). **NetLogo**. Center for Connected Learning and Computer-Based Modeling, Northwestern University. Evanston, IL. <http://ccl.northwestern.edu/netlogo>.

- 
- Wooldridge, M., & Jennings, N. R. (1995). **Intelligent agents: theory and practice**. *The Knowledge Engineering Review*, 10 (2), 115-152.