

تنظیم خودکار سختی بازی‌های توان‌بخشی با استفاده از روش یادگیری تقویتی چندتناوبی (یاقوت)

یونس سخاوت^۱، استادیار؛ حسین زارعی^۲، دانشجوی کارشناسی ارشد

۱- دانشکده چندرسانه‌ای - دانشگاه هنر اسلامی تبریز - تبریز - ایران - sekhavat@tabriziau.ac.ir

۲- دانشکده چندرسانه‌ای - دانشگاه هنر اسلامی تبریز - تبریز - ایران - hossein.zarei@tabriziau.ac.ir

چکیده: تحقیقات نشان داده است که طراحی فعالیت‌های توان‌بخشی در قالب بازی‌های جدی می‌تواند منجر به انگیزش بیشتر در بیماران شود. درجه سختی چنین بازی‌هایی معمولاً به‌صورت دستی توسط درمانگر تنظیم می‌شود. این در حالی است که بازی‌های توان‌بخشی خانگی نیاز به تنظیم خودکار درجه سختی متناسب با مهارت‌های بیماران دارند. این مقاله روشی برای تنظیم درجه سختی بازی‌های توان‌بخشی ارائه می‌کند که در آن سختی بازی به‌صورت خودکار و بر اساس مهارت‌های حرکتی بیماران و در حین انجام بازی تعیین می‌شود. این تحقیق نشان می‌دهد که تنظیم سختی بازی یک مسئله چندهدفی است که در آن اهداف مختلف ممکن است در دوره‌های تناوبی متفاوتی بررسی شوند. بدین منظور روشی تحت عنوان یادگیری تقویتی چند تناوبی (یاقوت) ارائه شده است که امکان ارزیابی اهداف مختلف را در دوره‌های تناوبی متفاوت فراهم می‌کند. نتایج آزمایش‌های انجام‌شده بر روی این سامانه حاکی از موفقیت قابل توجه آن در برآورده کردن معیارهای رضایت کاربران و همچنین بهبود مهارت‌های حرکت دست در مقایسه با روش‌های معمول یادگیری تقویتی است.

واژه‌های کلیدی: یادگیری تقویتی، تنظیم خودکار سختی بازی، دستگاه‌های توان‌بخشی، بازی‌های جدی

Dynamic Difficulty Adjustment of Rehabilitation Games using Reinforcement Learning

Y. A. Sekhvat¹, Assistant Professor; H. Zarei², MA Student

1- Faculty of Multimedia, Tabriz Islamic Art University, Tabriz, Iran, Email: sekhavat@tabriziau.ac.ir

2- Faculty of Multimedia, Tabriz Islamic Art University, Tabriz, Iran, Email: hossein.zarei@tabriziau.ac.ir

Abstract: Research has shown that the design of a rehabilitation task as a therapeutic game can result in a motivating rehabilitation environment. Generally, the difficulty level of a therapeutic game is regulated manually by a therapist. However, home-based rehabilitation games require a technique for automatic difficulty adjustment. This paper proposes a personalized difficulty adjustment technique for a rehabilitation game that automatically regulates difficulty settings based on a patient's skills in real-time. To this end, ideas from reinforcement learning are used to dynamically adjust the difficulty of a game. We show that difficulty adjustment is a multiple-objective problem, in which some objectives might be evaluated at different periods. To address this problem, we propose and use Multiple-Periodic Reinforcement Learning that makes it possible to evaluate different objectives of difficulty adjustment in separate periods. The results of experiments show that this technique outperforms traditional Multiple-Objective Reinforcement Learning in terms of user satisfaction parameters.

Keywords: Reinforcement learning (RL), dynamic difficulty adjustment, serious games, rehabilitation games.

تاریخ ارسال مقاله: ۱۳۹۵/۰۸/۲۴

تاریخ اصلاح مقاله: ۱۳۹۵/۱۰/۱۳

تاریخ پذیرش مقاله: ۱۳۹۵/۱۱/۲۰

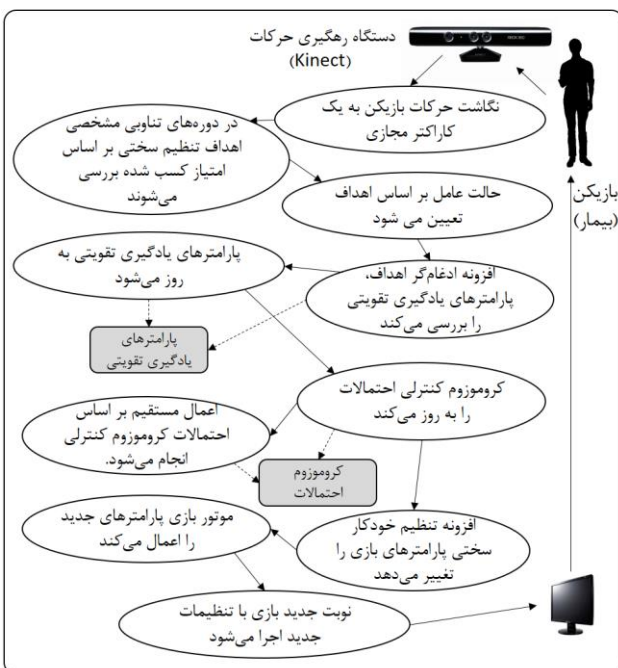
نام نویسنده مسئول: یونس سخاوت

نشانی نویسنده مسئول: ایران - تبریز - بلوار آزادی - میدان حکیم نظامی - دانشگاه هنر اسلامی تبریز - دانشکده چندرسانه‌ای.

۱- مقدمه

نوآوری اصلی این مقاله ارائه روشی برای تنظیم خودکار و شخصی‌سازی شده یک بازی توان‌بخشی است که در آن سختی بازی بر اساس پیشرفت بیمار تنظیم می‌شود. این سیستم با بهره‌گیری از فرآیند شناختی بدن، تعادلی بین درک بازیکن از محیط بازی و عمل او در حین انجام بازی برقرار می‌کند. فرآیند کلی روش ارائه‌شده در شکل ۱ نمایش داده شده است. برای پیاده‌سازی این سیستم، روشی تحت عنوان یادگیری تقویتی چندتناوبی (یاقوت) ارائه شده است که نسخه توسعه‌یافته یادگیری تقویتی چندمقصوده (یاقوم) [۱] است. مهم‌ترین ویژگی روش یاقوت امکان بررسی اهداف مختلف تنظیم سختی بازی در بازه‌های زمانی متفاوت است. این مقاله ادعا دارد که استفاده از این روش تأثیر قابل توجهی در بهبود عملکرد تنظیم درجه سختی بازی و به تبع آن ایجاد تجربه‌ای رضایت‌بخش در بازی دارد. نوآوری‌های این مقاله را می‌توان در موارد زیر خلاصه کرد:

- ارائه یک بازی توان‌بخشی برای بهبود حرکات دست که در آن درجه سختی بازی به صورت خودکار و بر اساس مهارت‌های حرکتی بازیکن تنظیم می‌شود.
- توسعه روش یادگیری تقویتی چندمقصوده (یاقوم) و ارائه روشی تحت عنوان یادگیری تقویتی چندتناوبی (یاقوت) برای فراهم کردن امکان ارزیابی اهداف مختلف در دوره‌های زمانی متفاوت.
- ارائه یک مدل احتمالی تنظیم سختی که امکان کشف شرایط جدید در عین بهره‌گیری از یادگیری‌های قبلی را ممکن می‌سازد.
- انجام مجموعه از آزمایش‌های میان‌مدت که تأثیر سیستم ارائه‌شده بر بهبود مهارت‌های حرکتی بیماران را بررسی می‌کند.



شکل ۱: شمایی کلی از فعالیتهایی که در روش یادگیری تقویتی چندتناوبی (یاقوت) انجام می‌شود.

انواع مختلفی از آسیب‌های مغزی از قبیل پارکینسون و سکتة مغزی می‌توانند منجر به اختلال در کنترل حرکت شوند [۱]. در این میان توان‌بخشی فیزیکی تأثیر مهمی در بازیابی توانایی‌های بیماران و درمان شناختی این افراد دارد. تمرکز توان‌بخشی بر روی ترمیم حرکت، هماهنگی و کنترل حرکت است که می‌تواند با تمرین‌هایی مانند حرکت عضو آسیب‌دیده و یا گرفتن یک شیء و کنترل آن انجام شود. علاوه بر کیفیت و تعداد حرکات انجام‌شده، شرکت فعال و باانگیزه بیماران در جلسات توان‌بخشی اهمیت ویژه‌ای در عملکرد درمان دارد [۲-۴]. این در حالی است که معمولاً فعالیت‌های توان‌بخشی در قالب کارهای تکراری و خسته‌کننده انجام می‌شوند که منجر به عدم رغبت بیماران برای شرکت در تمرین‌های توان‌بخشی می‌شوند. بر اساس تحقیقات انجام‌شده [۵]، بیمارانی که از فعالیت‌های توان‌بخشی لذت می‌برند زمان و انرژی بیشتری برای بهبود مهارت‌های خود صرف می‌کنند. در نتیجه، طراحی فعالیت‌های توان‌بخشی در قالب بازی (که به عنوان بازی‌های جدی شناخته می‌شوند)، می‌تواند منجر به افزایش انگیزه بیماران و به تبع آن افزایش اثربخشی این فعالیت‌ها شود [۶].

ویژگی مهم دیگر طراحی فعالیت‌های توان‌بخشی در قالب بازی‌های جدی، امکان ایجاد محیط شخصی‌سازی شده برای بیماران بر اساس توانایی‌های آن‌ها می‌باشد. این امر نیازمند شناخت ویژگی‌های یک بازی جدی و تأثیر این ویژگی‌ها بر روی درجه سختی بازی است. این امکان درمانگر را قادر می‌سازد تا سختی فعالیت‌ها را بر اساس پیشرفت بیمار تنظیم کند. تحقیقات نشان داده که زمانی که درجه سختی بازی متناسب با توانایی‌های یک بازیکن باشد، حس غوطه‌وری بیشتری در بازیکن ایجاد شده [۷-۱۲] و به تبع آن رضایت‌مندی بازیکن از بازی افزایش پیدا می‌کند.

در مراجع [۱۳، ۱۴] روشی برای تنظیم سختی ارائه شده است که در آن درجه سختی بازی در شروع بازی و بر اساس پروفایل کاربران تنظیم می‌شود. این درحالی است که مهارت‌های کاربران در طول بازی تغییر می‌یابند که نیازمند تغییر در پارامترهای سختی به صورت دستی توسط درمانگر می‌باشند. استفاده از قوانین رفتاری رویکرد دیگری برای تنظیم درجه سختی بازی است که در آن قوانین از پیش تعریف‌شده‌ای برای تنظیم درجه سختی مورد استفاده قرار می‌گیرند. برای مثال در مرجع [۱۵] قوانینی تعریف شده است که در آن زمانی که بازی کردن برای بازیکن سخت می‌شود، بازیکن سلاح‌های پیشرفته‌تری دریافت کرده و با دشمنان کم‌تری روبه‌رو می‌شود. با این حال شناسایی و پیاده‌سازی چنین قوانینی ساده نبوده و می‌تواند با اشتباهات زیادی همراه باشد. در مرجع [۱۶] روشی ارائه شده است که در آن فراخوانی هر یک از قوانین بر اساس احتمال نسبت داده‌شده برای آن قانون (که با توجه به نرخ موفقیت و شکست بازیکن تعیین می‌شود) انجام می‌شود. با این حال، شناخت و نگهداری جدول احتمالات با افزایش پیچیدگی‌های بازی غیرممکن خواهد بود.

۲- پیش‌زمینه

در این قسمت پس از بررسی تئوری‌های مربوط به تنظیم درجه سختی بازی، مفاهیم اولیه یادگیری تقویتی که مبنای روش ارائه شده در این مقاله است شرح داده شده است.

۲-۱- تنظیم درجه سختی بازی رایانه‌ای

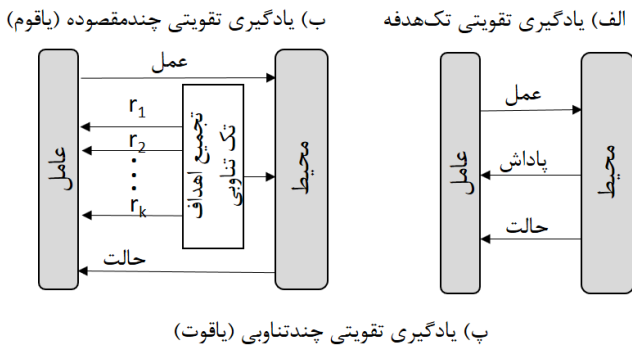
تنظیم درجه سختی بازی فرآیند پیدا کردن روش‌های مناسب برای ایجاد تعادل میان توانایی‌های فیزیکی افراد با حالت‌های موجود در یک بازی است. بر اساس تئوری روان بودن در بازی، بزرگی چالش‌های موجود در یک بازی وابسته به مهارت‌های بازیکن است [۱۷-۱۹]. بر اساس تئوری رابطه میان انگیزش و یادگیری، بیش‌ترین کارایی در یادگیری زمانی حاصل می‌شود که درجه سختی آن فعالیت در حد متوسط باشد. از این رو چالش‌ها و سختی بازی نباید به صورت خیلی ساده و یا خیلی سخت طراحی شوند. با توجه به اینکه عوامل و ویژگی‌های مختلفی در تنظیم سختی بازی تأثیر گذارند، تنظیم درجه سختی بازی‌های توان‌بخشی به صورت دستی حتی برای درمانگر نیز کار ساده‌ای نیست. علاوه بر آن، بازی‌های توان‌بخشی خانگی نیز نیاز به تنظیم خودکار درجه سختی بر اساس کارایی بیماران دارند. این سیستم خودکار می‌بایست در لحظه عمل کرده، به نحوی که اخلالی در روند بازی ایجاد نکند [۲۰، ۲۱].

تحقیقات قابل توجهی در زمینه تنظیم خودکار درجه سختی بازی‌های توان‌بخشی صورت گرفته است. در مرجع [۱۱]، یک بازی توان‌بخشی ارائه شده است که از فعالیت‌هایی چون حرکت دست و گرفتن اشیاء برای بهبود حرکات دست استفاده شده است. در این بازی انطباق سختی بازی با مهارت‌های بازیکن بر اساس بازخورد دریافتی از بازیکن انجام می‌شود. در مرجع [۲۲]، روشی برای تحلیل خودکار بازی‌های رایانه‌ای با استفاده از شبکه پتری رنگی ارائه شده است. در مرجع [۲۳]، چهارچوبی برای تنظیم سختی بازی توان‌بخشی ارائه شده است که در آن از بازخوردهای حیاتی (مانند ضربان قلب و مقدار تعریق بدن) استفاده شده است. این مقاله اعتقاد دارد که تنظیم درجه سختی بازی یک مسئله چندهدفه است که برخی از این اهداف ممکن است همسو و یا مخالف هم باشند. در نتیجه باید از روش جامعی برای در نظر گرفتن هم‌زمان این اهداف استفاده شود.

۲-۲- تنظیم درجه سختی با استفاده از یادگیری تقویتی

یادگیری تقویتی یک رویکرد محاسباتی برای خودکارسازی یادگیری هدفمند است [۲۴-۲۶]. این روش شامل فرآیند نگاشت حالت‌های یک عامل به کارهایی است که آن عامل می‌تواند بر روی محیط انجام دهد. در هر قدم، عامل اقدام به انتخاب و انجام بهترین عمل به صورتی می‌کند که میزان پاداش دریافتی در بلندمدت را افزایش دهد (شکل ۲-الف). در مسئله تعیین درجه سختی، امتیاز بازیکن به عنوان اصلی‌ترین معیار در تعیین حالت یک عامل است. فعالیت‌هایی که یک عامل برای تعیین سختی بازی انجام می‌دهد شامل ایجاد تغییراتی در ویژگی‌های بازی

متناسب با پارامترهای سختی بازی است (برای مثال تغییر در سرعت حرکت کاراکتر بازی). از طریق مشاهده تبعات تغییر در این پارامترهای سختی، عامل یاد می‌گیرد تا در آینده تصمیمات بهتری در جهت افزایش پاداش انجام دهد. استفاده از روش‌های یادگیری در طراحی ربات‌ها و کنترل کاراکترهای مستقل از اهمیت ویژه‌ای برخوردار است [۲۷]. در این میان، روش یادگیری تقویتی در کاربردهای مختلفی مورد استفاده قرار گرفته است. در مرجع [۲۸] روشی برای کنترل‌کننده تطبیقی بار فرکانس مبتنی بر یادگیری تقویتی ارائه شده است که نشان از توانایی این روش در تنظیم پویای پارامترهای سیستم با هر اندازه و پیچیدگی دارد.



شکل ۲: رابطه میان عامل و محیط و پاداش دریافتی در فرآیند یادگیری برای یادگیری تقویتی تک‌هدفه (الف) یادگیری تقویتی چندمقصوده (ب) و یادگیری چندتناوبی (پ)

در یادگیری تقویتی، متناسب با حالت فعلی (s) فعلی یک عامل، عمل a انتخاب و انجام شده که منجر به تغییر حالت عامل به حالت s' شده و پاداشی به میزان $r(s, a)$ دریافت می‌شود. در این روش، هدف عامل افزایش پاداش‌ها با یادگیری بهترین سیاست تصمیم‌گیری برای انتخاب یک عمل است. یک سیاست تصمیم‌گیری تابعی به صورت $a \rightarrow \pi(s)$ است که مجموعه‌ای از حالات یک عامل را به یک عمل نگاشت می‌کند. روش یادگیری Q [1] یکی از الگوریتم‌های پرکاربرد یادگیری تقویتی است که به صورت تکراری و قدم‌به‌قدم به محاسبه تابع ارزش-عمل بر اساس قانون زیر می‌پردازد:

$$Q(s, a) \leftarrow (1 - \alpha)Q(s, a) + \alpha(r + \gamma V(s') - Q(s, a)) \quad (1)$$

شاخص‌های مختلف در تناوب‌های متفاوت بررسی شوند. فاصله میان دو نوبت از بازی به‌عنوان دوره تناوب در نظر گرفته شده است که نسبت به نوع بازی تعیین می‌شود.

در مرحله شناسایی پارامترهای تاثیرگذار در دشواری بازی، کلیه پارامترهایی که با تغییر در آن‌ها می‌توان درجه سختی بازی را تغییر داد شناسایی می‌شوند. این پارامترها که به پارامترهای سختی معروف هستند، نسبت به نوع بازی تعیین می‌شوند. برای مثال، در بازی توان‌بخشی ارائه‌شده در این مقاله، سرعت حرکت کاراکتر، اندازه توپ و فاصله میان طاق‌های توپ به‌عنوان پارامترهای سختی تعریف شده‌اند.

مرحله نهایی، مشخص کردن الگوریتم و روش تغییر در مقادیر پارامترهای تاثیرگذار در دشواری بازی با توجه به نوع بازی است. در این مقاله از روش یادگیری تقویتی برای این منظور استفاده شده است که با توجه به نرخ یادگیری، مقدار جدید پارامترهای سختی در هر مرحله بر اساس مقدار قبلی پارامتر و مقدار پاداش جدید تعیین می‌شود. در ادامه این مقاله، پس از توضیح دو مورد ویژگی‌های بازی ارائه‌شده برای توان‌بخشی، جزئیات مربوط به پیاده‌سازی روش یادگیری تقویتی بر روی این بازی شرح داده شده است.

۴- نمونه موردی: بازی توان‌بخشی

بازی توان‌بخشی ارائه‌شده در این مقاله یک بازی مبتنی بر دستگاه کینکت است که از این ابزار برای رهگیری حرکات دست استفاده می‌کند. فضای بازی یک محیط سرسبز پوشیده‌شده با درختان و ساختمان‌های مختلف است. در این بازی حرکت فیزیکی دست‌های بازیکن به حرکات دست‌های یک کاراکتر مجازی نگاشت شده، به‌طوری‌که بازیکن کنترل دست‌های این کاراکتر مجازی را (که به‌صورت اول‌شخص نمایش داده می‌شود) بر عهده دارد. سناریوی این بازی توان‌بخشی شامل ضربه زدن به توپ‌هایی است که بر روی طاقی از توپ‌ها قرار دارد (شکل ۳ و شکل ۴).



شکل ۳: نمایی از بازی توان‌بخشی ارائه‌شده برای بهبود حرکت دست



شکل ۴: بازیکن در حال انجام بازی توان‌بخشی ارائه‌شده

در این تابع $V(s') = \max_a q(s, a)$ و α نرخ یادگیری است و γ به‌عنوان فاکتور تخفیف است که نشان‌دهنده اهمیت پاداش دریافت‌شده فعلی نسبت به پاداش‌های آتی است. روش‌ها و تئوری‌های یادگیری تقویتی به‌شکل وسیعی در سال‌های اخیر مورد استفاده قرار گرفته که منجر به پیشرفت‌هایی در برخی از مسائل جهان واقعی شده است [۱۲]. یادگیری تقویتی چندمقصوده (یا قوم) یکی از نسخه‌های گسترش‌یافته یادگیری تقویتی است که در آن چندین هدف مورد بررسی قرار می‌گیرند. همان‌طور که در شکل ۲-ب نمایش داده شده است، i بازخورد متناسب با هدف i می‌باشد. در صورتی‌که اهداف همسو باشند، امکان استخراج یک هدف با در نظر گرفتن توأمان اهداف وجود دارد. در صورتی‌که اهداف نامربوط باشند، توابع پاداش جداگانه‌ای می‌توانند مدنظر قرار گیرند. مشکل جایی ایجاد می‌شود که اهداف با یکدیگر تناقض داشته و نیاز به سبک و سنگین کردن اهداف باشد. این مقاله بر این باور است که مسئله تعیین سختی یک مسئله چندهدفه است که نیاز به در نظر گرفتن هم‌زمان اهداف دارد.

۳- کلیات روش ارائه شده

روشی که در این مقاله برای تنظیم درجه سختی بازی ارائه شده است، یک روش فراگیر است که می‌تواند برای تنظیم درجه سختی انواع بازی‌ها که در آن مجموعه‌ای از پارامترها تعیین‌کننده سختی بازی بوده و امتیاز بازیکن معیاری برای تنظیم این پارامترهاست، مورد استفاده قرار گیرد. با توجه به اینکه بازی‌های توان‌بخشی مثال بارزی از بازی‌هایی هستند که در آن تنظیم درجه سختی بازی بر اساس مهارت و عملکرد بیمار از اهمیت ویژه‌ای برخوردار است، یک بازی توان‌بخشی به‌عنوان نمونه موردی برای پیاده‌سازی روش یادگیری تقویتی چندتناوبی ارائه‌شده در این مقاله مورد استفاده قرار گرفته است.

به‌طورکلی، روش ارائه‌شده در این مقاله برای تنظیم درجه سختی بازی شامل مراحل (۱) شناسایی شاخص‌های توازن دشواری بازی (۲) تعیین تناوب بررسی شاخص‌های توازن دشواری (۳) شناسایی پارامترهای تاثیرگذار در دشواری بازی و (۴) مشخص کردن الگوریتم و روش تغییر در پارامترهای تاثیرگذار در دشواری بازی با توجه به نوع بازی می‌باشد. در مرحله شناسایی شاخص‌های توازن دشواری بازی، مواردی چون میزان رضایت بازیکن، چالش‌های بازی و میزان سرگرم‌کننده بودن بازی می‌تواند مورد توجه قرار گیرد. نکته مهم در این میان عددی کردن این شاخص‌ها برای تعیین میزان برآورده‌شدن آن‌ها با توجه به امتیاز کسب‌شده توسط بازیکن است. شاخص‌های (g_1, g_2, g_3) که در بخش ۱-۴ ارائه شده است، نمونه‌ای از شاخص‌های توازن دشواری هستند که بر اساس میزان رضایت بازیکن طراحی شده و برای بازی توان‌بخشی ارائه شده در این مقاله تعیین شده‌اند.

در مرحله تعیین تناوب بررسی شاخص‌های توازن دشواری، با توجه به اینکه شاخص‌های مختلف ممکن است در تناوب‌های یکسانی امکان بررسی نداشته باشند، روش یادگیری تقویتی این امکان را فراهم می‌کند که

استفاده قرار گیرند. حل این مسئله با استفاده از روش‌های موجود یادگیری تقویتی چندهدفه دارای محدودیت‌هایی می‌باشد. اول اینکه در الگوریتم‌های موجود، تمامی اهداف در دوره‌های تناوبی یکسانی مورد بررسی قرار می‌گیرند. درحالی‌که مسئله تنظیم درجه سختی بازی دارای اهدافی است که لزوماً در دوره‌های تناوبی یکسانی نمی‌توانند مورد بررسی قرار گیرند. برخی اهداف تنها بر اساس امتیاز پایان هر نوبت قابل بررسی است، درحالی‌که برای ارزیابی اهداف دیگر، تاریخچه‌ای از ارزیابی‌های قبلی مورد نیاز است. برای مثال گرچه امتیاز بازیکن در انتهای هر نوبت از بازی برای ارزیابی g_1 کافی است، ارزیابی شاخص g_2 نیازمند امتیاز نوبت فعلی و نوبت قبلی است. از سویی دیگر، ارزیابی شاخص g_3 نیازمند اطلاع از امتیازهای چندین نوبت متوالی می‌باشد. این مسئله نیازمند تغییر در روش یادگیری تقویتی می‌باشد که امکان ارزیابی شاخص‌های مختلف در دوره‌های زمانی متفاوت فراهم کند. مسئله دوم آن است که تعیین درجه سختی نیازمند روشی است که امکان برقراری ارتباط میان اهداف مختلف و حالت‌های مربوطه را فراهم کند (بدون توجه به اینکه اهداف در جهت هم، مخالف هم و یا بی‌ارتباط هستند). در نتیجه یک عامل با توجه به اهداف مختلف می‌تواند در حالت‌های مختلف باشد. در این مقاله مفهومی تحت عنوان یادگیری تقویتی چندتناوبی ارائه شده است که این دو مسئله ذکر شده را پوشش می‌دهد.

۵- یادگیری تقویتی چندتناوبی

روش یادگیری تقویتی چندتناوبی (یاقوت) ارائه شده در این مقاله برای تنظیم درجه سختی بازی دارای سه ویژگی منحصر به فرد است که آن را از روش‌های معمول یادگیری تقویتی متمایز می‌کند. اول اینکه یاقوت امکان بررسی اهداف مختلف را در دوره‌های تناوبی متفاوت فراهم می‌کند. دوم اینکه این روش امکان بررسی ارتباط میان حالت‌های مختلف با اهداف مختلف را فراهم می‌کند. سوم اینکه یاقوت یک مدل احتمالی دوسطحی ارائه می‌کند که امکان تغییر غیرمستقیم پارامترهای سختی را فراهم می‌کند. در ادامه هر یک از این ویژگی‌ها به تفصیل شرح داده شده است.

۵-۱- چندتناوبی

روش یاقوت این امکان را فراهم می‌کند که برخی اهداف مکررتر از اهداف دیگر بررسی شوند. اهدافی که در دوره‌های تناوبی کوتاهی بررسی می‌شوند، نیاز به تاریخچه کوتاهی از نتایج قبلی داشته، درحالی‌که ارزیابی برخی اهداف دیگر نیازمند تاریخچه بلندمدتی از نتایج عملکرد عامل است. معماری کلی یاقوت در شکل ۵ نمایش داده شده است. در این شکل، t_i بر روی محور x بیانگر نوبت i ام از بازی است که هر نوبت شامل عبور از ۱۰ طاق توپ است. فاصله میان دو نوبت از بازی به عنوان دوره تناوب نامیده شده است. سه نوع دوره تناوب با اندازه‌های یک، دو و شش متناسب با اهداف g_1 ، g_2 و g_3 در این بازی در نظر گرفته شده که در شکل ۵ نمایش داده شده است. خط چین عمودی بیانگر اهدافی است که در انتهای هر دوره تناوبی بررسی می‌شوند.

در این بازی، کاراکتر اول شخص که توسط بازیکن کنترل می‌شود بر روی یک مسیر از پیش تعیین شده حرکت داده می‌شود. در زمان‌های مشخصی، طاقی از توپ‌ها در مسیر ظاهر می‌شود که بازیکن می‌بایست با حرکت دست خود به توپ روشنی که در میان توپ‌های خاموش قرار دارد ضربه بزند. در این سناریو، همراه شدن فعالیت انجام شده توسط بازیکن با مشاهده دست‌های مجازی بر روی صفحه نمایش شرایطی ایجاد می‌کند که منجر به بازسازی حرکت دست بیمار می‌شود.

سختی بازی بر اساس مجموعه‌ای از پارامترها تعیین می‌شود که به پارامترهای سختی معروف هستند. در بازی ارائه شده در این مقاله، این پارامترها شامل سرعت حرکت کاراکتر بر روی مسیر تعیین شده (سرعت)، اندازه توپ‌های موجود بر روی طاق (اندازه) و فاصله بین طاق‌هایی است که بر روی مسیر ظاهر می‌شوند (فاصله)، می‌باشند. در نتیجه، عملکرد بازیکن تابعی از این پارامترهاست که تعیین‌کننده سختی بازی هستند.

۴-۱- تعیین درجه سختی بازی ارائه شده

گرچه بازیکنان حرفه‌ای بازی‌های رایانه‌ای از بازی‌هایی لذت می‌برند که غلبه بر آن‌ها دشوار است [۲۹] تازه‌کارها معمولاً از بازی‌هایی لذت می‌برند که در عین چالشی بودن بتوان بر آن‌ها فائق آمد [۳۰]. باور عمومی بر آن است که سختی بیش از اندازه یک بازی یا راحتی زیاد از حد اثر سوئی در جذابیت آن بازی و جذب مخاطب دارد [۳۱]. نکته مهم آن است که سختی بازی یک مفهوم نسبی است که بسته به توانایی‌های بازیکن متغیر است.

در بازی ارائه شده در این مقاله، ضربه به توپ درست منجر به برنده شدن (W) و ضربه به توپ اشتباه (و یا عدم ضربه) منجر به باخت (L) می‌شود. همان‌گونه که در مرجع [۳۲] اشاره شده است، جهت افزایش جذابیت بازی، اختلاف میان بردها W و باخت‌ها L می‌بایست کمینه شود. به‌طور کلی، در این مقاله سه شاخص برای اندازه‌گیری میزان رضایت بازیکنان در یک بازی در نظر گرفته شده که به شرح زیر می‌باشند:

- g_1 : برای مساوی نگه داشتن تعداد بردها و باخت‌ها در بازی، مقدار $W-L$ می‌بایست به حداقل برسد.
- g_2 : با در نظر گرفتن r_i به عنوان متوسط امتیازهای یک بازیکن در دو نوبت متوالی i و $i+1$ جهت جلوگیری از پیشرفت و یا پسرفت بزرگ در بازی، مقدار $|r_{i+1}-r_i|$ می‌بایست به حداقل برسد.
- g_3 : بازیکن می‌بایست شاهد پیشرفت در بازی بوده و در هر نوبت امتیاز بالاتری نسبت به نوبت قبل کسب کند. به این منظور تعداد پیشرفت‌ها در بازی (که برابر $\sum_{i=0}^n (r_{i+1}-r_i) > 0$ می‌باشد) می‌بایست به حداکثر مقدار خود برسد.

۴-۲- نیاز به یادگیری تقویتی چندتناوبی

شاخص‌های عنوان شده برای اندازه‌گیری میزان رضایت کاربران (g_1, g_2, g_3) می‌توانند به عنوان اهداف مسئله تنظیم درجه سختی بازی مورد

۵-۳- عمل دوسطحی احتمالی

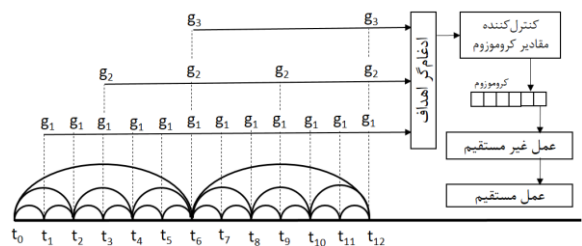
برخلاف روش‌های رایج یادگیری تقویتی که در آن عامل به صورت مستقیم و مشخص روی محیط عمل می‌کند، انجام یک عمل در روش یاقوت دوسطحی و احتمالی است. در این روش از دو نوع عمل مستقیم و غیرمستقیم استفاده می‌شود. عمل غیرمستقیم مسئول افزایش و یا کاهش احتمال انجام یک عمل مستقیم است. درحالی‌که عمل مستقیم بر روی محیط انجام می‌شود که می‌تواند منجر به تغییر حالت عامل شود. در بازی ارائه شده در این مقاله، اعمال مستقیم به صورت تغییر در پارامترهای سختی (شامل سرعت اندازه و فاصله) تعریف شده‌اند. یک کروموزوم کنترلی مطابق آنچه در شکل ۶ نمایش داده شده است، برای مشخص کردن احتمال تغییر در پارامترها استفاده شده است. این کروموزوم کنترلی شامل شش عدد حقیقی مرتبط با پارامترهای سختی است. همان‌گونه که در شکل ۶-الف نمایش داده شده است، هر عدد حقیقی بیانگر احتمال فعال کردن یک رفتار متناسب با پارامترهای سختی است. دو عدد اول مربوط به سرعت می‌باشد (P1 برای افزایش سرعت و P2 برای کاهش سرعت)، دو عدد بعدی مربوط به اندازه (P3 برای اندازه توپ و P4 برای کاهش اندازه توپ) و دو عدد آخر مربوط به فاصله می‌باشد (P5 برای افزایش فاصله و P6 برای کاهش فاصله). هر عدد بیانگر احتمال فعال کردن کنترل‌کننده (کاهش/افزایش) یکی از پارامترهای سختی است. در دوره‌های زمانی مشخص، کنترل‌کننده‌های افزایشی و یا کاهشی باعث تغییر در احتمال انجام یک عمل می‌شوند. عدد ۱ نشانگر فعال شدن و ۰- نشانگر غیرفعال شدن انجام یک عمل مستقیم است (شکل ۶-ب).

در آغاز بازی، کروموزوم کنترلی به صورت تصادفی مقدار می‌گیرد. در انتهای دوره‌های زمانی مختلف، کروموزوم بر اساس قوانینی که در الگوریتم ۱ نشان داده شده است (که به عنوان سیاست انتخاب عمل شناخته می‌شود) به‌روزرسانی می‌شود. با در نظر گرفتن t به عنوان شماره نوبت بازی، ارزیابی هدف g_i در دوره‌های زمانی P_i به نحوی که $t \bmod p_i = 0$ باشد صورت می‌گیرد. پس از ارزیابی یک هدف (و یا مجموعه‌ای از اهداف) در آن دوره زمانی، کروموزوم کنترلی C به‌روزرسانی می‌شود. در این کروموزوم c_i بیانگر مکان i ام در کروموزوم است که نشان‌دهنده احتمال فعال‌سازی (غیرفعال‌سازی) افزایش (کاهش) پارامترهای سختی (سرعت، فاصله و اندازه) است. همچنین $sign(p)$ نشان‌دهنده حالت کروموزوم قبل از به‌روزرسانی است. در صورتی که قبلاً فعال باشد، $sign(p)$ برابر ۱ و در صورتی که غیرفعال بوده باشد، $sign(p)$ برابر ۰- خواهد بود. پارامتر a بیانگر نرخ یادگیری برای یادگیری تقویتی است. نرخ یادگیری مشخص‌کننده آن است که الگوریتم به چه مقدار می‌بایست شرایط جدید را امتحان کرده و به چه میزان از یادگیری‌های قبلی استفاده کند. در صورت فعال‌سازی افزایش یک پارامتر، مقدار آن پارامتر ۱۰٪ افزایش پیدا کرده و در صورت فعال‌سازی کاهش، مقدار آن ۱۰٪ کاهش پیدا می‌کند. در روش یاقوت به جای به‌روزرسانی مقادیر $q(s,a)$ و انتخاب عمل مناسب بر اساس حالت موجود،

هدف g_1 در زمان‌های t_1, t_2, \dots هدف g_2 در زمان‌های t_2, t_4, \dots و هدف g_3 در زمان‌های t_6, t_{12}, \dots بررسی می‌شود. در صورت وجود بیش از یک هدف در انتهای هر یک از دوره‌های تناوبی، قسمتی تحت عنوان «ادغام‌گر اهداف» وجود دارد که مسئول تجمیع پارامترهای مربوط به اهداف مختلف است. با در نظر گرفتن k هدف g_1, \dots, g_k و شدت اهمیت متناسب با هر هدف $\lambda_1, \dots, \lambda_k$ ، پاداش تجمیع‌شده به روشی که در الگوریتم ۱ نشان داده شده محاسبه می‌شود.

۵-۲- چندحالتی

در روش یاقوت، هر هدف به مجموعه‌ای از حالات نسبت داده می‌شود، به نحوی که یک عامل یاقوت می‌تواند به صورت هم‌زمان در حالت‌های مختلفی نسبت به اهداف متفاوت باشد. با توجه به شاخصی رضایت بازیکن که در فصل قبل شرح داده شد، حالت‌های موجود به شرح زیر تعریف می‌شوند. برای g_1 مقدار $W-L$ مشخص‌کننده حالت عامل یاقوت در پایان هر نوبت از بازی (عبور از ۱۰ طاق توپ) است. با توجه به اینکه $W+L=10$ است، مجموعه حالت‌های بازیکن با توجه به g_1 شامل $\{-10, -9, \dots, 9, 10\}$ خواهد بود. در مورد g_2 که اختلاف میان دو نوبت متوالی مدنظر است، مقدار $|r_{i+1}-r_i|$ مشخص‌کننده حالت بازیکن است. با توجه به اینکه امتیاز هر نوبت عددی بین ۰ و ۱۰ است، حالت‌های ممکن برای g_2 شامل مجموعه $s_2 = \{-10, -9, \dots, 9, 10\}$ می‌باشد. برای مشخص کردن حالت عامل نسبت به هدف g_3 نتایج شش نوبت متوالی به عنوان تعیین‌کننده حالت بازیکن نسبت به g_3 در نظر گرفته شده است. با توجه به تابع $\sum_{i=0}^n (r_{i+1} - r_i > 0)$ به عنوان خروجی برای هدف g_3 حالات ممکن شامل $s_3 = \{5, 4, 3, 2, 1\}$ خواهد بود. در حالت کلی، یک عامل یاقوت با سه تایی $\langle s_1, s_2, s_3 \rangle$ شناخته می‌شود که $s_i \in S_i$ می‌باشد.



شکل ۵: معماری یادگیری تقویتی چندتناوبی (یاقوت) که در آن اهداف مختلف در دوره‌های تناوبی متفاوت بررسی می‌شوند.

	Speed	Size	Distance				
	↑+	↓-	↑+	↓-			
(الف)	P1	P2	P3	P4	P5	P6	← کروموزوم احتمالات بر اساس پارامترهای سختی
(ب)	0.24	0.85	0.54	0.08	0.47	0.13	← نمونه ای از کروموزوم با مقادیر احتمال انجام یک عمل
(پ)	-1	1	1	-1	1	-1	← فعال یا غیرفعال شدن انجام یک عمل بر اساس احتمالات

شکل ۶: کروموزوم کنترلی برای مشخص کردن احتمال انجام یک عمل (افزایش یا کاهش پارامترهای سختی)

شرکت‌کننده به مدت ۸ جلسه در طول دو هفته (هر روز حداکثر یک جلسه) از بازی توان‌بخشی استفاده کرد. در هر جلسه درمانی از شرکت‌کنندگان خواسته شد تا به مدت ۲۰ دقیقه به صورت آزاد با بازی توان‌بخشی ارائه‌شده بازی کنند. جلسات آزمایش در محل آزمایشگاه واقعیت افزوده شناختی واقع در دانشگاه هنر اسلامی تبریز انجام شد. تعداد ۸ نفر از شرکت‌کنندگان از سیستم تنظیم درجه سختی یاقوت و ۸ نفر دیگر از سیستم یاقوم استفاده کردند.

از یک رویکرد احتمالی استفاده شده است که در آن تمامی اعمال مستقیم بر اساس احتمال نسبت داده شده به آن عمل شانس اجرا دارند. در الگوریتم یاقوت، هر هدف اثر خود را بر روی پاداش تجمعی (rw) اعمال می‌کند. پارامتر λ_i متناسب با هدف g_i مشخص‌کننده اهمیت و اولویت آن هدف در محاسبه پاداش تجمعی است. در پایان الگوریتم و پس از به‌روزرسانی مقادیر کروموزوم کنترلی، عمل مستقیم افزایش و یا کاهش در پارامترهای سختی انجام می‌شود.

۶- آزمایش‌ها و ارزیابی

نمونه اولیه بازی توان‌بخشی توصیف‌شده در این مقاله به همراه افزونه تنظیم درجه سختی بازی در موتور بازی‌سازی یونیتی (Unity 5.3.4) پیاده‌سازی شده و از مازول *Kinect Developer Toolkit 1.8.0* و همچنین *KinectSDK v.1.8.0* برای اتصال بازی به دستگاه کینکت استفاده شده است. کلید اسکرپت‌های مربوط به تنظیم درجه سختی بازی و همچنین کنترل کاراکتر بازی با زبان *C#* پیاده‌سازی شده که از تارنمای آزمایشگاه واقعیت افزوده شناختی www.carlab.ir دانشگاه هنر اسلامی تبریز قابل دریافت است. ابتدا آزمایش‌هایی انجام شد تا مقدار نرخ یادگیری با توجه به نوع مسئله تنظیم درجه سختی به صورت تجربی تعیین شود. سپس آزمایش‌هایی بر روی ۱۶ شرکت‌کننده انجام شد تا اثر روش ارائه‌شده بر روی بهبود مهارت‌های حرکتی دست و رضایت افراد بررسی شود. کلید افراد مورد آزمایش از بین افراد سالم انتخاب شده و با انجام یک سری آزمایش‌ها این اطمینان حاصل شد که شرکت‌کنندگان دارای مشکلات روانی و شناختی نیستند.

۶-۱- مطالعه اثر تغییر نرخ یادگیری

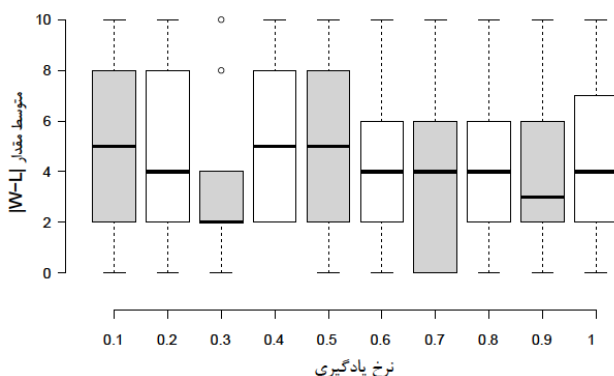
به منظور مطالعه اثر تغییر نرخ یادگیری α بر عملکرد بازیکنان و به تبع آن تعیین مناسب نرخ یادگیری، آزمایش‌هایی ترتیب داده شد که در آن نرخ یادگیری با اندازه‌های ۰/۱ تا ۱ (با فاصله یک‌دهم) مورد بررسی قرار گرفت. با در نظر گرفتن g_i به عنوان شاخص اصلی رضایت، هدف بررسی این مطلب بود که چه نرخ تغییری موجب کم‌ترین مقدار $|W-L|$ می‌شود. اثر تغییر پارامتر α بر روی پارامتر رضایت کاربر در شکل ۷ نمایش داده شده است. گرچه روند مشخصی با تغییر مقدار α قابل مشاهده نیست، مقدار ۰/۳ برای α منجر به بهترین نتیجه شده است. از این رو این مقدار در آزمایش‌های بعدی به عنوان نرخ یادگیری استفاده شده است.

۶-۲- مطالعه کارایی روش یاقوت

به منظور مطالعه و ارزیابی کارایی روش تقویتی چندتناوبی (یاقوت) و مقایسه آن با روش یادگیری تقویتی چندمقصوده (یاقوم) نوع سیستم به عنوان متغیر مستقل در این مطالعه مورد توجه قرار گرفت. هدف اصلی این آزمایش‌ها مقایسه عملکرد تنظیم درجه سختی در بهبود مهارت‌های حرکتی دست و همچنین رضایت‌مندی شرکت‌کنندگان از انجام بازی بود. بدین منظور یک مطالعه میان‌مدت طراحی شد که در آن هر

```
repeat
  Let  $G = null$  (goals that are evaluated at the current
  round).
  float  $rw = 0$ ; (representing aggregatedRewards)
  for  $i=1$  to  $k$  do
    if  $t \bmod p_i == 0$  then
       $G = G \cup g_i$ 
  end
  foreach  $g_i$  in  $G$  do
    if  $g_i == g_1$  then
       $rw_1 = s_1 + (|t.score| - 5)$ 
      aggregatedRewards +=  $g_1 * \lambda_i * rw_1$ ;
    if  $g_i == g_2$  then
       $rw_2 = s_2 + (|t.score| - |(t-1).score|)$ 
      aggregatedRewards +=  $g_2 * \lambda_i * rw_2$ ;
    if  $g_i == g_3$  then
       $rw_3 = s_3 + \sum_{i=t}^{t-5} (r_{i+1} - r_i > 0)$ 
      aggregatedRewards +=  $g_3 * \lambda_i * rw_3$ ;
  end
   $rnd = \text{Random}(0,1)$ ;
  if ( $rnd \leq 0.3$ ) then
    for  $p = 1$  to 6 do
      if ( $p$  is even) then
         $C_p = (1 - \alpha) * C_p + \alpha * (sign(p) + rw)$ 
      if ( $p$  is odd) then
         $C_p = (1 - \alpha) * C_p - \alpha * (sign(p) + rw)$ 
      end
    end
    for  $p = 1$  to 6 do
      with the probability of  $C_i$  in  $C$ , run the
      corresponding direct action
      update the states  $\{s_1, s_2, s_3\}$ 
    end
  end
until the game is progress;
```

الگوریتم ۱: تغییر در مقادیر پارامترهای تأثیرگذار در دشواری بازی

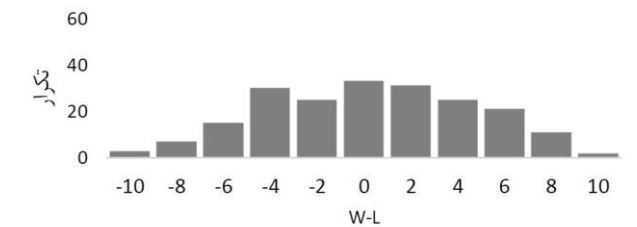


شکل ۷: اثر تغییر نرخ یادگیری بر روی میزان رضایت‌مندی بازیکن

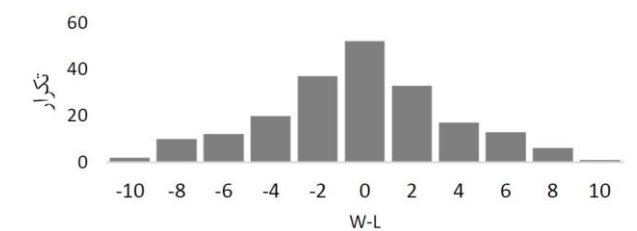
۶-۲-۱- شاخص احساس رضایت از انجام بازی

آزمایش‌هایی ترتیب داده شد تا مشخص شود که یاقوت و یاقوم تا چه میزان می‌توانند در برآورده کردن شاخص احساس رضایت با توجه به اهداف g_1, g_2, g_3 موفق باشند. بر اساس g_1 کم کردن مقدار $|W-L|$ منجر به احساس رضایت بیشتر در بازیکن و به تبع آن سرگرم شدن بیشتر می‌شود. با توجه به هیستوگرام اختلاف میان تعداد بردها و باخت‌ها که در شکل ۸ نمایش داده شده است، در مقدار قابل‌توجهی از موارد در روش یاقوت مقدار $|W-L|$ صفر و یا نزدیک به صفر است. به‌طور مشخص، در بیش‌تر از ۸۰٪ موارد اختلاف بین امتیازها ۴ یا کم‌تر از ۴ است که نشان‌دهنده سرگرم‌کننده بودن بازی در بیش‌تر از ۸۰٪ موارد است. این در حالی است که بر اساس هیستوگرام روش یاقوم، میانگین مقدار $|W-L|$ نزدیک صفر نبوده و نمودار حالت تقارنی ندارد. این مشاهدات حاکی از عملکرد بهتر روش یاقوت در برآورده کردن g_1 است.

هیستوگرام یاقوم



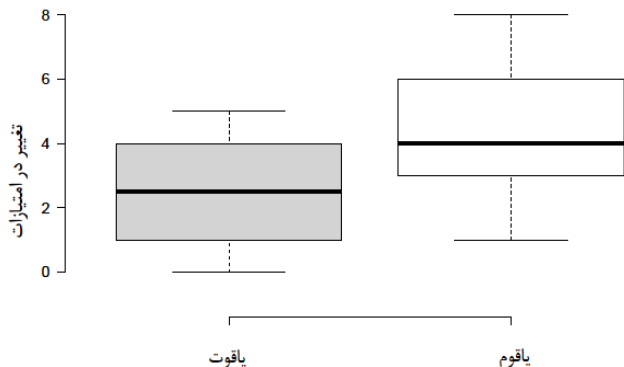
هیستوگرام یاقوت



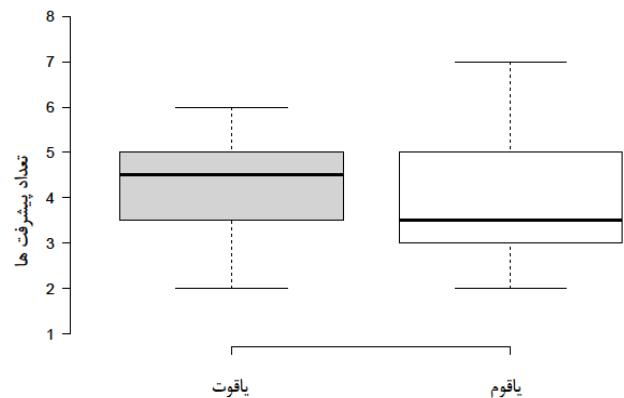
شکل ۸: هیستوگرام تفاوت میان بردها و باخت‌ها $|W-L|$ در دو روش یاقوم و یاقوت

عملکرد روش یاقوت در مقایسه با روش یاقوم در کم کردن تغییرات در نوبت‌های متوالی (در ارتباط با g_2) و همچنین بیشینه کردن تعداد پیشرفت‌ها (در ارتباط با g_3) مورد بررسی قرار گرفت. با توجه به نتایجی که در شکل ۹ نمایش داده شده است، متوسط مقدار تغییر در امتیازات روش یاقوم است. بر اساس آزمون تی انجام‌شده بر روی داده‌های کاربران اختلاف آماری معناداری (سطح معناداری ۰/۰۲۶۳۱) میان مقدار تغییر در امتیازات در روش یاقوت (میانگین ۲/۵۶۲۵ و انحراف استاندارد ۱/۵۷۹۹) و یاقوم (میانگین ۴/۶۲۵ و انحراف استاندارد ۱/۷۸۰۹) وجود دارد. در مورد بیشینه کردن میانگین تعداد پیشرفت‌ها در نوبت‌های متوالی ($\sum_{i=0}^n (r_{i+1} - r_i) > 0$) همان‌طور که در شکل ۱۰ قابل مشاهده است، روش‌های یاقوت و یاقوم رفتارهای نسبتاً مشابهی از خود نشان داده‌اند. نتیجه آزمون تی نیز حاکی عدم وجود اختلاف آماری معنادار (سطح معناداری ۰/۷۳۶) میان تعداد پیشرفت‌ها در روش یاقوت

(میانگین ۴/۱۵۶ و انحراف معیار استاندارد ۱/۰۹۲) و یاقوم (میانگین ۳/۸۱۲ و انحراف معیار استاندارد ۱/۴۴۵) است. این امر می‌تواند به‌سادگی بازی توان‌بخشی ارائه‌شده نسبت داده شود که امکان پیشرفت برای بازیکن در هر دو روش یاقوم و یاقوت به‌سادگی میسر است.



شکل ۹: مقایسه روش یاقوت و یاقوم با توجه به شاخص g_2 (مقدار تغییر در امتیازات در نوبت‌های مختلف)



شکل ۱۰: مقایسه روش یاقوت و یاقوم با توجه به شاخص g_3 (تعداد پیشرفت‌ها در نوبت‌های مختلف)

۶-۲-۲- بهبود مهارت‌های حرکتی

در روش یادگیری تقویتی چند تناوبی که در این مقاله ارائه شده است، درجه سختی بازی بر اساس توانایی‌های حرکتی بازیکنان تنظیم می‌شود. در نتیجه افزایش درجه سختی در طول جلسات بازی بیانگر بهبود در مهارت‌های حرکتی بازیکنان است؛ اما با توجه به اینکه بازیکنان مختلف مهارت‌های مختلفی در شروع آزمایش‌ها دارند، به‌جای مقایسه مقدار پارامترهای سختی، میزان تغییر در این پارامترها به‌عنوان معیاری از بهبود در توانایی‌های حرکتی مورد توجه قرار گرفته است. بدین منظور، میزان تغییر در میانگین پارامترهای سختی بازی (سرعت، اندازه و فاصله) به‌عنوان پارامترهای بهبود در مهارت‌ها مورد توجه قرار گرفت. همان‌گونه که در شکل ۱۱ و ۱۲ نشان داده شده است، نرخ کاهش در اندازه توپ‌ها و همچنین فاصله میان طاقه‌ای توپ برای روش یاقوت بیشتر از روش یاقوم است. این مشاهدات بیانگر افزایش شدت حرکت دست‌ها در یاقوت در مقایسه با روش یاقوم است که عامل مهمی در توان‌بخشی حرکتی است. در مورد عامل سرعت، گرچه مقدار سرعت در طول ۸ جلسه برای

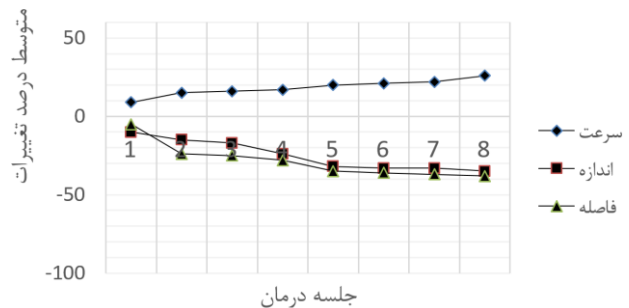
آزمایش‌ها حاکی از بهبود بیشتر در عملکرد حرکتی شرکت‌کنندگان با استفاده از روش یاقوت در مقایسه با روش یاقوم است. برای مطالعات آتی، نویسندگان مقاله قصد دارند تا با کشف روابط میان اهداف مختلف مسئله تنظیم درجه سختی، مدلی جامع‌تری از روش یادگیری چندتناوبی ارائه کنند.

مراجع

- [1] N. Hocine, A. Gouaich, and S. A. Cerri, "Dynamic difficulty adaptation in serious games for motor rehabilitation," *International Conference on Serious Games*, pp. 115-128, Berlin, Germany, 2014.
- [2] K. D. O. Andrade, T. B. Pasqual, G. A. Caurin, and M. K. Crocorno, "Dynamic difficulty adjustment with evolutionary algorithm in games for rehabilitation robotics," *IEEE International Conference on Serious Games and Applications for Health*, pp. 1-8, Orlando, USA, 2016.
- [3] M. C. Cirstea, and M. F. Levin, "Improvement of arm movement patterns and endpoint control depends on type of feedback during practice in stroke survivors," *Neurorehabilitation and Neural Repair*, vol. 21, pp. 398-411, 2007.
- [4] B. H. Dobkin, "Rehabilitation after stroke," *New England Journal of Medicine*, vol. 352, no. 16, pp. 1677-1684, 2005.
- [5] N. Hocine, and A. Gouaich, "Therapeutic games' difficulty adaptation: An approach based on player's ability and motivation," *International Conference on Computer Games Conference*, pp. 257-261, Louisville, USA, 2011.
- [6] D. Afegan, E. M. Peck, E. T. Solovey, A. Jenkins, S. W. Hincks, E. T. Brown, and R. J. Jacob, "Dynamic difficulty using brain metrics of workload," *ACM Conference on Human Factors in Computing Systems*, pp. 3797-3806, Toronto, Canada, 2014.
- [7] Y. A. Sekhavat, and P. Abdollahi, "Can google nowcast the market trend of Iranian mobile games?," *In International Conference on Web Research*, Tehran, Iran, 2016.
- [8] Y. A. Sekhavat, "KioskAR: an augmented reality game as a new business model to present artworks," *International Journal of Computer Games Technology*, vol. 2106, pp. 1-12, 2016.
- [9] L. F. Lucca, "Virtual reality and motor rehabilitation of the upper limb after stroke: a generation of progress?," *Journal of Rehabilitation Medicine*, vol. 41 no. 12, pp. 1003-1006, 2009.
- [10] A. Gouaich, N. Hocine, L. Van Dokkum, and D. Mottet, "Digital-pheromone based difficulty adaptation in post-stroke therapeutic games," *ACM SIGHIT International Health Informatics Symposium*, pp. 5-12, Miami, USA, 2012.
- [11] A. Heuser, H. Kourtev, S. Winter, D. Fensterheim, G. Burdea, V. Hentz, and P. Fordeucey, "Telerehabilitation using the Rutgers Master II glove following carpal tunnel release surgery: proof-of-concept," *Transactions on Neural Systems and Rehabilitation Engineering*, vol. 15 no. 1, pp. 43-49, 2007.
- [12] C. Liu, X. Xu, and D. Hu, "Multiobjective reinforcement learning: A comprehensive overview," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 45 no. 3, pp. 385-398, 2015.
- [13] S. Natkin, C. Yan, S. Jumpertz, and B. Market, "Creating multiplayer ubiquitous games using an adaptive narration model based on a user's model," *Digital Games Research Association International Conference*, Tokyo, Japan, 2007.
- [14] G. N. Yannakakis, and J. Hallam, "Real-time adaptation of augmented-reality games for optimizing player satisfaction," *IEEE Symposium On Computational Intelligence and Games*, pp. 103-110, Perth, Australia, 2008.
- [15] R. Hunnicke, and V. Chapman, "AI for dynamic difficulty adjustment in games," *Challenges in Game AI Workshop, Nineteenth National Conference on Artificial Intelligence*, San Jose, USA, 2004.
- [16] P. Spronck, I. Sprinkhuizen-Kuyper, and E. Postma, "Difficulty scaling of game AI," *International Conference on Intelligent Games and Simulation*, pp. 33-37, Ghent, Belgium, 2004.
- [17] S. Prabhakaran, E. Zarahn, C. Riley, A. Speizer, J. Y. Chong, R. Lazar, and J. W. Krakauer, "Inter-individual variability in the

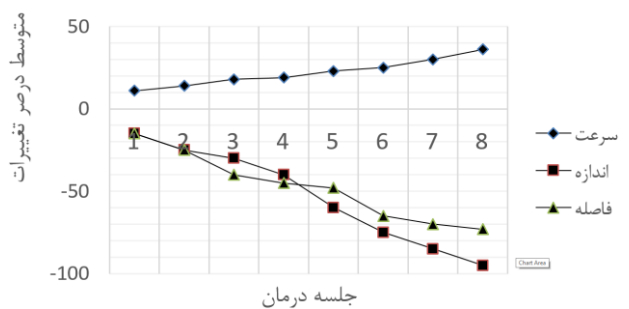
هر دو روش افزایش پیدا کرده است، بهبود قابل توجهی بر روی عامل سرعت مشاهده نشده است. توجیه این مشاهده نیازمند کشف روابط میان پارامترهای سختی است که در مطالعات آتی مورد توجه قرار خواهد گرفت.

یاقوم



شکل ۱۱: متوسط درصد تغییرات برای پارامترهای سختی بازی ارائه شده در روش یاقوم

یاقوت



شکل ۱۲: متوسط درصد تغییرات برای پارامترهای سختی بازی ارائه شده در روش یاقوت

۷- نتیجه‌گیری

در این مقاله یک بازی توان‌بخشی برای بهبود مهارت‌های حرکت دست با افزونه‌ای برای تنظیم خودکار درجه سختی ارائه شد. پس از شناسایی پارامترهای سختی بازی، سه هدف و شاخص که بیان‌کننده میزان رضایت کاربران از بازی هستند بر اساس تئوری‌های مختلف تعیین شدند. در این مقاله نشان داده شد که ماهیت این اهداف به‌گونه‌ای است که در دوره‌های تناوبی یکسانی قابل بررسی نیستند. برخی از این اهداف صرفاً بر اساس امتیاز نهایی در هر نوبت قابل بررسی بوده و برخی دیگر نیاز به تاریخچه‌ای از امتیازات نوبت‌های قبلی دارند. برای حل این مشکل روشی تحت عنوان یادگیری تقویتی چندتناوبی (یاقوت) ارائه شد که نسخه توسعه‌یافته روش یادگیری تقویتی چندمقصوده (یاقوم) است. در ادامه روش یاقوت از لحاظ ساختاری و معنایی شرح داده شده و روشی دوسطحی برای انجام اعمال یک عامل تعریف شد. نتایج آزمایش‌های میان‌مدت بر روی تعدادی از کاربران نشان داد که روش یاقوت عملکرد بهتری در برآورده کردن شاخص‌های جلب رضایت کاربران در مقایسه با روش یاقوم داشته است. همچنین، نتایج حاصل از

- [25] H. Wang, Y. Gao, and X. Chen, "RI-dot: A reinforcement learning npc team for playing domination games," *IEEE Transactions on Computational Intelligence*, vol. 2, no. 1, pp. 17-26, 2010.
- [26] R. S. Sutton, and A. G. Barto, *Reinforcement learning: An introduction*, Cambridge: MIT press, 1998.
- [۲۷] عادل اکبری مجد، حسین شایقی، حمید محمدنژاد، عبدالله یونسی، «کنترل‌کننده مقاوم تطبیقی بار فرکانس مبتنی بر یادگیری تقویتی برای یک سیستم قدرت به‌هم‌پیوسته شامل SMES»، *مجله مهندسی برق دانشگاه تبریز*، دوره ۴۷، شماره ۲، صفحات ۳۸۱-۳۹۰، ۱۳۹۶.
- [۲۸] علیرضا رضائی، ابوالقاسم اسدالله راعی، سعید شیری قیداری، «یادگیری رفتار مقاوم در مقابل تغییرات محیطی و خرابی حسگرهای روبات سیار، با استفاده از شبکه بیزین پویای مبتنی بر داده»، *مجله مهندسی برق دانشگاه تبریز*، دوره ۴۳، شماره ۱، صفحه ۲۷-۳۸، تابستان ۱۳۹۲.
- [29] M. Buro, and T. Furtak, "RTS games as test-bed for real-time AI research," *Joint Conference on Information Science*, pp. 481-484, Cary, USA 2003.
- [30] B. Scott, "The illusion of intelligence," *AI Game Programming Wisdom*, vol. 1 no.1, pp. 16-20, 2002.
- [31] T. W. Malone, "What makes things fun to learn? Heuristics for designing instructional computer games," *ACM symposium and the first SIGPC symposium on Small systems*, pp 162-169, Palo Alto, USA 1980.
- [32] J. Hagelback, and S. J. Johansson, "Measuring player experience on runtime dynamic difficulty scaling in an RTS game," *IEEE Symposium on Computational Intelligence and Games*, pp. 46-52, Milan, Italy, 2009.
- capacity for motor recovery after ischemic stroke," *Neurorehabilitation and Neural Repair*, 2007.
- [18] M. Csikszentmihalyi, *Flow: the psychology of optimal Experience* Harper Collins Berkely, California: Osborne/McGraw-Hill, 1990.
- [19] C. H. Tan, K. C. Tan, and A. Tay, "Dynamic game difficulty scaling using adaptive behavior-based AI," *IEEE Transactions on Computational Intelligence*, vol. 3, no. 4, pp.289-301, 2011.
- [20] M. S. Cameirão, S. B. i Badia, E. D. Oller, and P. F. Verschure, "Neurorehabilitation using the virtual reality based Rehabilitation Gaming System: methodology, design, psychometrics, usability and validation," *Journal of Neuroengineering and Rehabilitation*, vol. 7, no. 1, pp. 1, 2007.
- [21] G. Andrade, G. Ramalho, H. Santana, and V. Corruble, "Extending reinforcement learning to provide dynamic game balancing," *Workshop on Reasoning, Representation, and Learning in Computer Games, 19th International Joint Conference on Artificial Intelligence*, pp. 7-12, York, NY, USA, 2005.
- [۲۲] سعید پاشازاده، «تحلیل خودکار بازی رایانه‌ای با استفاده از شبکه پتری رنگی»، *مجله مهندسی برق دانشگاه تبریز*، دوره ۴۶، شماره ۲، صفحات ۳۷-۴۸، تابستان ۱۳۹۵.
- [23] A. Gouaïch, N. Hocine, L. Van Dokkum, D. Mottet, "Digital-pheromone based difficulty adaptation in post-stroke therapeutic games," *ACM SIGHIT International Health Informatics Symposium*, pp. 5-12, Miami, FL, USA, 2012.
- [24] Y. Chen, W. Xu, H. Sundaram, T. Rikakis, and S. M. Liu, "A dynamic decision network framework for online media adaptation in stroke rehabilitation," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 5, no. 1, pp.4, 2008.