

## تعیین خودکار توابع تمایز برای دسته‌بندی الگوها

محمد رضامیبدی

غلامرضا ضابطی

دانشکده مهندسی کامپیوتر

دانشگاه صنعتی امیرکبیر

تهران - ایران

### چکیده

مسئله یادگیری قاعده تصمیم‌گیری جهت دسته‌بندی الگوهای با نظارت<sup>(۱)</sup> با هدف مینیمم کردن تعداد دسته‌بندی‌های غلط، یکی از مسائل قدیمی محسوب می‌شود، یکی از روش‌های دسته‌بندی الگوها، استفاده از تابع تمایز<sup>(۲)</sup> می‌باشد، در این مقاله روش‌های جدیدی براساس بازی آناتامونهای یادگیر جهت تعیین تابع تمایز ارائه شده است، در این روشها، تابع تمایز از طبقه یک جستجوی احتمالات<sup>(۳)</sup> در فضای پارامترهای تابع تمایز بدست می‌آید، آنچه ما در این مقاله بر آن تاکید داریم، استفاده از آناتامونهای یادگیر با ساختار ثابت به عنوان بازیگران بازی و بررسی کارایی این آناتامونها در یادگیری تابع تمایز بهینه خطی با غیر خطی می‌باشد.

**کلمات کلیدی:** شناسایی الگو، تابع تمایز، آناتامونهای یادگیر

### ۱- مقدمه

این مقاله به مسئله دسته‌بندی الگوهای با نظارت با هدف مینیمم کردن تعداد دسته‌بندی‌های غلط می‌پردازد، اگر  $P(w_i|x)$  احتمال شرطی تعلق داشتن بردار  $x$  به کلاس  $w_i$  باشد، با استفاده از قانون:

$$x \in w_i \text{ if } P(w_i|x) > P(w_j|x), j \neq i$$

بهترین قاعده تصمیم‌گیری که تعداد دسته‌بندی‌های غلط را مینیمم می‌کند، حاصل می‌شود<sup>[۱][۲]</sup> [۱۳] چون غالباً توزیع  $P(w_i|x)$  معلوم نیست، این قانون باید با استفاده از یک سری نمونه‌های آموزشی با دسته‌بندی شناخته شده از الگوها، یادگرفته شود، برای یادگیری این قاعده تصمیم‌گیری، دو رویکرد کلی وجود دارد، یک رویکرد، تخمین تابع چگالی شرطی  $f(x|w_i)$  از روی مجموعه نمونه‌ای آموزشی می‌باشد که در این صورت، با داشتن احتمال هر کلاس  $P(w_i)$ ، قانون بجز<sup>(۴)</sup>، می‌تواند قاعده تصمیم‌گیری را تعیین کند<sup>[۱]</sup>.

رویکرد دیگر استفاده از تابع تمایز است، در یک مسئله دو کلاسه، اگر<sup>(۵)</sup>  $g(x) > 0$ ، تابع تمایز باشد، قاعده، تصمیم‌گیری به صورت زیر خواهد بود:

$$\begin{cases} x \in w_i & \text{if } g(x) > 0 \\ x \in w_j & \text{if } g(x) < 0 \end{cases}$$

روشهای متعددی جهت تعیین تابع تمایز ارائه شده است، در صورتیکه کلاس‌های الگوها به طور خطی جداش پذیر باشند، با استفاده از الگوریتم‌های پرسریزی، می‌توان تابع تمایز به فرم  $g(X) = \theta^T X$  (۶) علامت ترتیبه داشت<sup>(۷)</sup> را تعیین کرد<sup>[۱][۲]</sup> [۱۳]. در صورتیکه کلاسها به طور خطی جایابی پذیر باشند، روش Ho-Kashaplike با استفاده از مینیمم کردن متوسط خطای منبعات، تابع تمایز را تعیین می‌کند<sup>[۱۳]</sup> [۱]. برخی از تکنیکهای یادگیری تابع تمایز را تعیین می‌کنند<sup>(۸)</sup> [۱۳] [۳]. در این روشها، هدف، پیدا کردن بردار وزن  $\theta$  به گونه‌ای است که  $\theta^T X = g(X)$  مجموع خطای

منبعات را مینیمم کند و یا به نوع تقریب از  $(w_i|x)$  را حاصل کند، هیچ کدام از روش‌های فوق الزاماً احتساب دسته‌بندی‌های غلط را مینیمم نمی‌کند، زیرا برآوردن شرطی مانند مینیمم کردن متوسط خطای منبعات، الزاماً به معنی مینیمم کردن احتمال دسته‌بندی غلط نیست.

روشهای بیز براساس آناتامونهای یادگیر<sup>(۹)</sup> جهت تعیین تابع تمایز ارائه شده است، Barto Andrew روش نام  $A_{R,p}$  ارائه کرده‌اند که هدف آن شناسایی تابع تمایز<sup>(۱۰)</sup>  $g(X) = \theta^T X$  با استفاده از آناتامونهای یادگیر می‌باشد<sup>[۳]</sup> [۱۳]. در این روش یک ربطه لجیس<sup>(۱۱)</sup> این هر عرض از نمونه ورودی و هر عمل تصمیم‌گیری ایجاد می‌شود و آناتامونهای یادگیر سعی می‌کنند که با مشاهده هر نمونه ورودی، تصمیم صحیح را تاختاذ کنند، در این روش از آناتامونهای یادگیر با ساختار متغیر با دو عمل استفاده شده و از الگوریتم شیوه الگوریتم Robins - Monro<sup>(۱۲)</sup> [۳] با هدف مینیمم کردن متوسط خطای منبعات استفاده شده است، این روش محدودیت خطی پودن تابع تمایز را دارد و همچنین به دلیل برقراری یک ربطه لجیس<sup>(۱۳)</sup> بروزی مجموعه نمونه‌ای آموزشی، مجموعه نمونه‌ها باید محدود باشند.

Sastry و Thatachar<sup>(۱۴)</sup> روش را جهت یادگیری تابع تمایز بهینه از فضای تابع تمایزها استفاده از بازی آناتامونهای یادگیر با پاسخ پیکان ارائه کرده‌اند<sup>[۱۳]</sup> [۱]. الگوریتم ارائه شده در این روش بر اساس آناتامونهای یادگیر با ساختار متغیر می‌باشد که احتمال هر عمل با استفاده از تمحیص از ماتریس پاداش<sup>(۱۵)</sup> بازی که در طول یادگیری ساخته می‌شود، به عکام می‌شود، هر آناتامون شرکت کننده در بازی، یکی از پارامترهای تابع تمایز را تعیین می‌کند، مجموعه اعمال هر آناتامون، مجموعه مقادیر ممکن برای آن پارامتر می‌باشد، در این روش، کلیه ترکیهای ممکن عملیات آناتامونهای شرکت کننده در بازی به مظور تعیین ماتریس پاداش بازی، در حافظه ذخیره می‌شود، بازی ارائه

Bays Rule<sup>(۱۶)</sup>  
Pay Off Matrix<sup>(۱۷)</sup>

Probabilistic Search<sup>(۱۸)</sup>  
Associative Relation<sup>(۱۹)</sup>

Discriminate Functions<sup>(۲۰)</sup>  
Learning Automata<sup>(۲۱)</sup>

Supervised Pattern Recognition<sup>(۲۲)</sup>  
Stochastic Approximations<sup>(۲۳)</sup>

نه در این روش، ساس کار این مقاله را تشکیل می‌دهد با این تفاوت که ما از آناتوئهای پادگیر با اختصار ثابت استفاده می‌کنیم.

$\Phi = \{\Phi_1, \dots, \Phi_S\}$  مجموعه حالات داریم آناتوئون،  $\underline{\alpha} = \{\alpha_1, \dots, \alpha_T\}$  مجموعه عملها یا خروجی‌های آناتوئون که بینگر مجموعه ورودی آناتوئون است.  $(\cdot, \cdot, \cdot)$  تابع تغییر حالت آناتوئون بر اساس حالت ورودی فعلی است، یعنی:

$$F(\cdot, \cdot) : \Phi \times \underline{\alpha} \rightarrow \Phi$$

$$\Phi(k+1) = F[\Phi(k), \beta(k)]$$

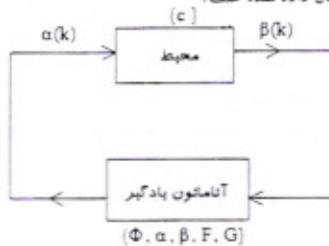
تابع خروجی آناتوئون می‌باشد که بر اساس حالت فعلی، خروجی آناتوئون را تعیین می‌کند، یعنی:

$$G(\cdot) : \Phi \rightarrow \underline{\alpha} : \alpha(k) = G[\Phi(k)]$$

محیط بوسیله ۳ تابی  $(\underline{\alpha}, \underline{c}, \underline{\beta})$  تعریف می‌شود که در آن  $\underline{\alpha} = \{\alpha_1, \dots, \alpha_T\}$  مجموعه ورودی‌ها،  $\underline{\beta} = \{\beta_1, \dots, \beta_m\}$  مجموعه خروجی‌های محیط و  $\underline{c} = \{c_1, \dots, c_T\}$  مجموعه مشخصات داخلی محیط است بطوریکه هر  $c_i$  به صورت زیر ربطه دارد:

$$c_i = \Pr[\beta(k) = \alpha_i]$$

۴ امثال دریافت پاسخ نامطلوب از جانب محیط است اگر عمل  $\alpha_i$  در محیط تمام گیرد. آناتوئون پادگیر سعی می‌کند تا با استفاده از پاسخهایی که از جانب محیط دریافت می‌کند، عملی را که با نجام آن، احتمال دریافت پاسخ نامطلوب از جانب محیط مینبینم گردد را پیدا کند. ربطه محیط و آناتوئون که بصورت پیشخور می‌باشد در شکل (۱) شناس داده شده است.



شکل ۱: رابطه آناتوئون پادگیر با محیط

از نظر ساختاری، آناتوئهای پادگیر به دسته تقسیم می‌شوند: آناتوئهای پادگیر با اختصار ثابت (۸) و آناتوئهای پادگیر با اختصار متغیر (۹). در آناتوئهای با اختصار ثابت، تابع تغییر حالت آناتوئون بوسیله تابع  $F(\cdot, \cdot, \cdot)$  از قبل مشخص می‌شود. بر اساس حالت آناتوئون در مرحله  $k$  ام، عمل  $\alpha(k)$  با استفاده از تابع  $G(\cdot)$  به محیط اعمال می‌شود و با دریافت پاسخ  $\beta(k)$  از جانب محیط حالت بعدی آناتوئون تعیین می‌گردد. بعض از مدل‌های آناتوئهای آناتوئون با اختصار ثابت در بخش ۷ این مقاله بررسی شده‌اند.

در آناتوئهای پادگیر با اختصار متغیر، تابع تغییر حالت از قبیل مشخص نیست. در این مدل، بردار احتمال عمل (۱۰) پا بردار احتمال حالت (۱۱) آناتوئون مورد توجه است. بردار احتمال عمل با حالت، احتمال انتخاب یک عمل با قرار گرفتن آناتوئون در یک حالت را در مرحله  $k$  ام تعیین می‌کند. من توان در حل مسائل این دو بردار را با حفظ شرایط معادل گرفت (۹). بردار احتمال عمل به صورت زیر تعریف می‌شود:

$$P(k) = \{P_1(k), \dots, P_T(k)\}$$

$$P_i(k) = \Pr[\alpha(k) = \alpha_i | \beta(k-1), \dots, \beta(1)]$$

در آناتوئهای پادگیر با اختصار متغیر، یک توزیع اولیه (ممولاً یکنواخت) به عنوان مقادیر اولیه بردار احتمال عمل در نظر گرفته می‌شود. در هر مرحله، آناتوئون یک عمل را با توجه به احتمالهای عملها، انتخاب و به محیط اعلام

نه در این روش، ساس کار این مقاله را تشکیل می‌دهد با این تفاوت که ما از آناتوئهای پادگیر با اختصار ثابت استفاده می‌کنیم.

روش را براساس شبکه‌ای پیشخور (۱۲) Phananskar و Thatachar استفاده از شبکه‌های سمعتی محدوده‌های هر کلاس تعیین زده می‌شود. در این روش از آناتوئهای پادگیر با اختصار متغیر با شیوه به عهده سازی  $L_{R-I}$  استفاده و نشان داده شده که می‌بینیم من تواند به یک نقطه بهینه محلی (۱۳) (و نه از اسرازی) اگر شرایط اولیه مناسب باشد، همگراشود.

روش را با ارائه PLA (Parameterized PLA) و Sastry (Parameterized LA) (۱۴) جهت شناسایی نقطه بهینه سراسری (۱۵) با استفاده از شبکه‌های پیشخور پیشنهاد کردند [۸] [۷]. در این روش بحای آناتوئهای استاندارد از آناتوئون جدیدی به PLA که دارای ساختار متغیر بوده و قابلیت دریافت بردار محبوی (۱۶) از محیط را دارد، استفاده و نشان داده شده است که تحت شرایط بیزیه‌ای، مجموعه آناتوئها می‌تواند به نقطه بهینه سراسری همگراشود.

تجهیز در کنیه روش‌های ارائه شده قابل بر اساس آناتوئهای پادگیر شناسی است. استفاده از آناتوئهای پادگیر با اختصار متغیر به عنوان اجزای بازی پا به استفاده پادگیری می‌باشد. در این مقاله برای پادگیری تابع تغییر بجای استفاده از آناتوئهای پادگیر با ساختار متغیر، از آناتوئهای پادگیر با ساختار ثابت استفاده گردیده و نشان داده شده است که آناتوئهای با ساختار ثابت، کارایی بالاتری داشتند که آناتوئهای پادگیر با ساختار متغیر دارای نباشد. بخش‌های بعدی مقاله به صورت نیز تطبیق شده است.

در بخش ۲ به معرفی اجمالی آناتوئهای پادگیر می‌پردازیم و در بخش ۳ بازی آناتوئهای پادگیر با پاسخ یکسان را معرفی می‌کنیم. در بخش ۴ یک مسئله تئوریک سه کلاسه را تعریف می‌کنیم که در طول مقاله، روش‌های مختلف روی آن پیاده شود. در بخش ۵ روش ارائه شده توسط Sastry و Thatachar بررسی می‌شود. در بخش ۶ به بررسی آناتوئهای پادگیر با ساختار متغیر استاندارد و در بخش ۷ به بررسی آناتوئهای پادگیر با ساختار ثابت جدیدی که دارای کارایی بالاتری هستند پیشنهاد می‌گردد و در بخش ۸ به نتیجه‌گیری می‌پردازیم.

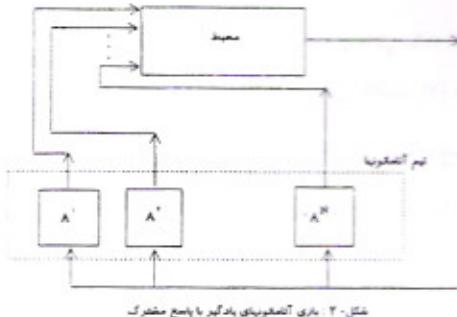
## ۲ - آناتوئون پادگیر

ایده آناتوئهای پادگیر به مدل‌سازی رفتار تصمیم‌گیری می‌شوند که نویزیکی بازی می‌گردد. این مدل، بر اساس انتخاب یک عمل از مجموعه محدود از اعمال با توجه به تجربیات گذشته، بنائده است. آناتوئون پادگیر در یک محیط تصادفی ناشناس فعالیت می‌کند، عمل را به محیط اعلام می‌کند و پاسخ آن را از محیط دریافت می‌نماید و براساس پاسخ دریافت شده سعی در پادگیری عمل بهینه دارد.

یک آناتوئون پادگیر، یک آناتوئون احتمالی (۱۷) است که با محیط خود به صورت پیشخور تراکنش دارد. خروجی آناتوئون عملی است که به عنوان ورودی، به محیط داده می‌شود و خروجی محیط پاسخ محیط به عمل انجام شده است که به عنوان ورودی به آناتوئون ارسال می‌گردد. آناتوئون پادگیر براساس فعالیت آناتوئون، انجام اعمالی از مجموعه اعمال آناتوئون است که بهترین پاسخ را از محیط دریافت کند.

به طور رسمی، آناتوئون پادگیر بوسیله ۵ تابی

Global Optimum Point (۱۸)	Local Optimum Point (۱۹)	3-Layer Linear Piecewise Classification (۲۰)	Feed Forward Network (۲۱)
Feed Back (۲۲)	Stochastic Automaton (۲۳)	Context Vector (۲۴)	
Variable Structure Learning Automata (۲۵)	Fixed Structure Learning Automata (۲۶)		
	State Probability Vector (۲۷)	Action Probability Vector (۲۸)	



مقادیر ممکن برای پارامتر  $\theta$  (و در نتیجه اعمال آناتمانون  $A^1$ ) از قبل تعیین شده‌اند. هدف از این بازی، پیدا کردن مقادیر ممکن برای پارامترهای  $\theta$  می‌باشد. طوری که تعداد مستنبتدیهای خلف می‌بینیم گردد. در هر مرحله، آناتمانون  $A^1$  با توجه به شیوه تخصیم گیری خود، یک عمل (مقدار) از مجموعه اعمال خود (مجموعه مقادیر ممکن برای  $\theta$ ) را انتخاب می‌کند و به محیط اعلام می‌نماید. یک (جند) نمونه از بردارهای ویژگی موجود در محیط نوسط تابع تمايز بدلست آمده (با استفاده از پارامترهای انتخاب شده) مستنبتدی می‌شود و تیجه بصورت نسبت تعداد مستنبتدیهای صحیح به کل نمونهای آزمایش شده، به تهم ارسل می‌شود. براساس پاسخ بدست آمده، آناتمانوهای شرکت کننده در بازی، شیوه تخصیم گیری خود را به عنوان می‌نمایند. هدف از بازی، بدست اوردن تابع تمايزی است که بر اساس آن احتمال دریافت پاسخ نامطلوب و یا به عبارت دیگر تعداد دسته بندی‌های خلف نمونه‌ها نمینمی‌بینیم گردد.

#### ۴ - یک مسئله نمونه

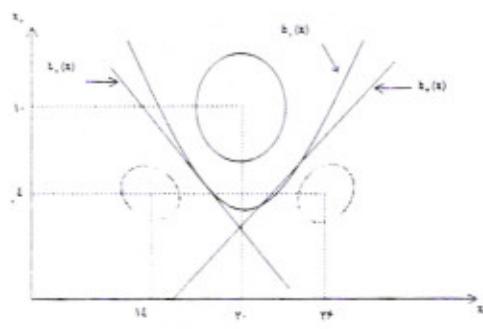
کلیه روش‌های مختلف پاده‌سازی شده در این مقاله، روی یک مسئله نمونه انجام شده است که در این بخش معرفی می‌گردد.  
مسئله: پیدا کردن تابع تمايز بینه برای سه کلاس  $W_1$ ,  $W_2$  و  $W_3$  با توزیعهای نرمال با مشخصات زیر:

$$W_1 : M = (20, 10)^T, \Sigma = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}$$

$$W_2 : M = (14, 4)^T, \Sigma = \begin{pmatrix} 2 & 0.5 \\ 0.5 & 1 \end{pmatrix}$$

$$W_3 : M = (26, 4)^T, \Sigma = \begin{pmatrix} 2 & 0.5 \\ 0.5 & 1 \end{pmatrix}$$

بردار  $M$ ، بردار مبانکن و ماتریس  $\Sigma$ ، ماتریس کوواریانس توزیعها می‌باشد. در شکل ۳، شکل توزیع کلاسها و تابع تمايز هر کلاس نشان داده شده است.



من نماید و پاسخ آن را از محیط دریافت می‌کند. براساس پاسخ دریافت شده آناتمانون، این احتمالها را به هنگام می‌نماید و در مرحله بعدی براساس احتمالهای بهعنوان شده، عمل بعدی را انتخاب می‌کند. هدف از این بهعنوان سازی، یافتن عمل مناسب با هدف دریافت پیشترین احتمال پاسخ مطلوب از محیط است. به نحوه تغییر احتمالها الگوریتم پادگیری<sup>(۱)</sup> یا روش شوین<sup>(۲)</sup> گفته می‌شود که می‌تواند تابع خطی یا غیرخطی از احتمالها باشد. برخی از این الگوریتم‌های پادگیری در بخش ۶ این مقاله بررسی شده‌اند.

آناتمانوهای پادگیر با ساختار ثابت براساس تابع  $(\cdot, \cdot, \cdot, \cdot)$  به عرضه قطعی<sup>(۳)</sup> و غیرقطعی<sup>(۴)</sup> تقسیم می‌شوند. در مدل قطعی، براساس حالت و ورودی فعلی آناتمانون، به طور قطعی می‌توان خروجی فعلی و حالت بعدی آناتمانون را تعیین کرد اما در مدل غیرقطعی، یکی از توابع (یا هردو)  $(\cdot, \cdot, \cdot, \cdot)$  غیرقطعی می‌باشد.

محیط براساس مجموعه  $\beta$  (مجموعه پاسخ محیط) به سه دسته تقسیم می‌شود:

- مدل P: در این حالت مجموعه  $\beta$  دو عضوی است که یک عضویانگر پاسخ مطلوب و عضو دیگر بیانگر پاسخ نامطلوب می‌باشد و عموماً به صورت  $\{\beta_1, \beta_2\}$  بیان می‌شود.

- مدل Q: در این مدل مجموعه خروجی به صورت مجموعه محدود  $\{\beta_1, \dots, \beta_m\}$  تعریف می‌شود. خروجی محیط می‌تواند یکی از مقادیر گسته مدل  $\beta_1, \dots, \beta_m$  را اختیار کند.

- مدل S: در این مدل مجموعه خروجی، یک مجموعه نامحدود به صورت  $\{\beta(a, b)\}$  تعریف می‌شود که  $a, b$  اعداد حقیقی هستند. پاسخ محیط می‌تواند یکی از مقادیر پیوسته در ناحیه  $(a, b)$  را اختیار کند. در مدل‌های S, Q, عموماً مقدار پیشتر برای پاسخ محیط به عنوان پاسخ بهتر در نظر گرفته می‌شود.

براساس مجموعه  $\beta$ ، محیط به درسته است<sup>(۵)</sup> و پس از طبقه‌بندی می‌شود. در مدل استا مقدار  $\beta$ ‌ها و در نتیجه مشخصات محیط در طول زمان ثابت است در صورتیکه در مدل بپایان مشخصات داخلی محیط در طول پادگیری در حال تغییر می‌باشد.

برای اطلاعات بیشتر درباره آناتمانوهای پادگیر، می‌توانید به [۱۰][۱۱][۱۲][۱۳][۱۴] مراجعه کنید.

#### ۳ - بازی آناتمانها

از نظر تئوری، یک مسئله را می‌توان با یک آناتمانون متعدد که با محیط در ارتباط است حل کرد اما در این صورت ممکن است مجبور به استفاده از آناتمانون با تعداد عملهای بسیار زیاد بشویم که می‌تواند مجرّد برای سرعت همگرایی بسیار پایین گردد<sup>[۹]</sup>. برای حل این مشکل، می‌توان از مجموعه‌ای از آناتمانون که به صورت یک تهم در حل مسئله شرکت می‌کند، استفاده کرد. مدل‌های مختلفی و بازی آناتمانون از این شده است. در مدلی که در این مقاله در نظر گرفته شده است، همه بازیکران در انتهای هر بازی پاسخ یکسانی را از محیط دریافت می‌کنند<sup>[۹]</sup>.

همانطور که قبلاً اشاره شد، هدف از این مقاله، شناسایی پارامترهای تابع تمايز از طریق آناتمانوهای پادگیر می‌باشد. فرض کنیم تابع تمايز به فرم  $h(x) = g(\theta_1, \dots, \theta_N, x)$  باشد. طوری که  $\theta_1, \dots, \theta_N$  پارامترهای تابع  $g$  هستند که پاید یادگرفته شوند و  $x$  بردار نمونه ویژگی‌ها است که پاید دستنبتد شود. با استفاده از  $N$  آناتمانون پادگیر در قالب یک بازی با پاسخ یکسان مطابق شکل (۲) می‌توان پارامترهای فوق را تخمین زد.

طبقه‌بندی شکل (۲)، آناتمانون  $A^1, \dots, A^N$  در بازی شرکت کرده‌اند. هر آناتمانون  $A^i$  دارای  $T$  عمل  $\alpha_{1,i}, \dots, \alpha_{T,i}$  می‌باشد. هر عمل آناتمانون  $A^i$  یک مقدار ممکن برای پارامتر  $\theta$  از تابع تمايز در نظر گرفته می‌شود.

توسط ماتریس  $A_{\frac{n}{n}}$  که دارای  $\frac{n}{n}$  عمل است، انتخاب شده باشد، بردار احتمال عمل ماتریس توسط ربطه (۲) بعنوان می‌شود، در ربطه (۲)،  $E_{\frac{n}{n}}(k)$  پیشین احتمال دریافت پاسخ مطلوب از طرف محیط است اگر آناماتون  $A_{\frac{n}{n}}$  عمل  $j$  را در مرحله  $k$  ام انتخاب کرده باشد،  $E_{\frac{n}{n}}(k)$  با استفاده از ماتریس پاداش بازی در طول پادگیری ساخته می‌شود.تابع  $S_{\frac{n}{n}}(k)$  ملاکس جهت شریف با تابع کردن آناماتون  $A_{\frac{n}{n}}$  است که طبق رابطه (۳) تعریف می‌شود. تابع (۳) یک تابع اکیدا صعودی بین  $[0, 1]$  می‌باشد و  $\lambda_n$  پارامتر پادگیری است. با توجه به تحلیل الگوریتم در [۴] با انتخاب  $\lambda_n$  به حد کافی کوچک، نیم آناماتونها با احتمال زدیک به یک، به سمت لستراتیزی بهینه همگرا می‌شود.

$$\begin{aligned} P_{\frac{n}{n}}(k) &= P_{\frac{n}{n}}(k) - \lambda_n \left[ f\left(E_{\frac{n}{n}}(k)\right) - f\left(E_{\frac{n}{n}}^*(k)\right) \right] \\ &\quad \left[ S_{\frac{n}{n}}(k) P_{\frac{n}{n}}(k) + S_{\frac{n}{n}}(k) \left(1 - P_{\frac{n}{n}}(k)\right) \frac{P_{\frac{n}{n}}^*(k)}{r_{n-1}} \right], j \neq i \\ P_{\frac{n}{n}}^{(k+1)} &= 1 - \sum_{i=1}^{n-1} P_{\frac{n}{n}}^{(k+1)} \quad (۴) \\ S_{\frac{n}{n}}(k) &= 1 \text{ if } E_{\frac{n}{n}}(k) > E_{\frac{n}{n}}^*(k) \quad (۵) \\ &= 0 \text{ otherwise} \end{aligned}$$

الگوریتم (۱) سکلت برنامه پادگیری را برای پادگرفتن یک تابع تمايز نشان می‌دهد. این سکلت در کلیه روش‌های پیاده‌سازی شده در این مقاله بکسان است. در این الگوریتم، روش GenerateSamples جمجمه نمونه‌های آموزش و تست را تولید می‌کند. روش SelectActions عمل هر آناماتون را انتخاب می‌کند. روش Discriminate بوسیله عملیات انتخاب شده یک (یا چند) نمونه گرفته می‌شود. پس از انجام پادگیری، تست و گزارش کارایی روی جمجمه‌های ۵۰۰ تابع که متفاوت با جمجمه آموزش هستند، تجام می‌گیرد. در هر آزمایش، تابع اعلام شد، میانگین ۱۰۰ شبیه‌سازی می‌باشد.

مقادیر مسکن برای پارامترهای تابع تمايز مثله نمونه کا در ربطه (۱) آمده‌اند، براساس مشخصات توزیع نمونه‌های مسکن می‌باشند. انتخاب شده‌اند. اما در عمل، بدليل اینکه توزیع نمونه‌ها مشخص نیست، چنین امکانی وجود ندارد. برای رفع این بازی بزرگی برای هر پارامتر در نظر گرفت را با در نظر گرفتن یک ساختار درختی از آناماتونها پادگیر برای پادگیری هر پارامتر، پارامتر مورد نظر را تعیین کرد. در این ساختار سطوح بالاتر (زدیک به ریشه) یا نکر بازدهی بزرگ و سطوح پایه‌ی، یا نکر بازدهی کوچک هستند و برگها یا نکر مقادیر پارامترها می‌باشند. با توجه به دقت موره انتظار برای هر پارامتر، می‌توان تعداد سطوح ساختار درختی را انتخاب کرد. مقدار هر پارامتر با توجه به مسیر فعلی شده در هر درخت تعیین می‌گردد [۲].

تابع  $(x) = h_i(x)$ ، تابع تمايزی هستند که باید پادگرفته شوند پطریکه:

$$h_i(x) > 0 \Leftrightarrow x \in W_i$$

$$h_i(x) < 0 \Leftrightarrow x \notin W_i$$

با توجه به شکل توزیع، تابع:

$$h_i(x) = x_1 - 0.14x_1^2 + 0.08x_1 - 0.0 > 0 : x \in W_i$$

$$h_i(x) = -x_1 - x_1 + 0.23 > 0 : x \in W_i$$

$$h_i(x) = -x_1 + x_1 - 0.17 > 0 : x \in W_i$$

می‌توانند تغییرهای خوبی از تابع تمايز کلاس‌های مسئله پاشند.

همانطور که مشاهده می‌شود،  $(x)$  دارای ۳ پارامتر و  $h_i(x)$  دارای ۲ پارامتر هستند که باید پادگرفته شوند. بنابراین از سه تیم آناماتون مسکن شامل ۳، ۲ و ۲ آناماتون که به طور مستقل کار می‌کنند، برای پادگیری تابع تمايز استفاده می‌کنند:

$$h_1 : A_{11}, A_{12}, A_{13} \rightarrow h_1 = x_1 - \alpha_{11}x_1^2 + \alpha_{12}x_1 + \alpha_{13}$$

$$h_2 : A_{21}, A_{22} \rightarrow h_2 = -x_1 - \alpha_{21}x_1 + \alpha_{22}$$

$$h_3 : A_{31}, A_{32} \rightarrow h_3 = -x_1 + \alpha_{31}x_1 - \alpha_{32}$$

$i_j$  ها آناماتونهای پادگیری هستند که پارامترهای تابع تمايز  $(x)$  را تعیین می‌کنند.  $i_j$  ها پارامترهای تابع تمايز  $(x)$  هستند. مجموعه مقادیر مسکن برای پارامترهای این سه تابع تمايز بصورت زیر در نظر گرفته شده‌اند. مقادیر مسکن برای هر پارامتر، در جلوی نام آناماتون تعیین کننده آن پارامتر، آمده است.

$$A_{11}: 0/15, 0/17, 0/19, 0/21, 0/23 \quad (1)$$

$$A_{12}: 6/6, 7/1, 7/6, 8/2, 8/8 \quad A_{13}: 7/0, 7/5, 8/0, 8/5, 9/0$$

$$A_{21}: 0/0, 0/70, 1, 1/25, 1/0 \quad A_{22}: 1/9, 2/1, 2/2, 2/5, 2/7$$

$$A_{31}: 0/0, 0/70, 1, 1/25, 1/0 \quad A_{32}: 1/9, 2/1, 2/2, 2/5, 2/7$$

در قسمت پادگیری، برای هر کلاس ۵۰۰ نمونه تصادفی آموزش در نظر گرفته می‌شود. پس از انجام پادگیری، تست و گزارش کارایی روی جمجمه‌های ۵۰۰ تابع که متفاوت با جمجمه آموزش هستند، تجام می‌گیرد. در هر آزمایش، تابع اعلام شد، میانگین ۱۰۰ شبیه‌سازی می‌باشد.

مقادیر مسکن برای پارامترهای تابع تمايز مثله نمونه کا در ربطه (۱) آمده‌اند، براساس مشخصات توزیع نمونه‌های مسکن می‌باشند. انتخاب شده‌اند. اما در عمل، بدليل اینکه توزیع نمونه‌ها مشخص نیست، چنین امکانی وجود ندارد. برای رفع این بازی بزرگی برای هر پارامتر در نظر گرفت را با در نظر گرفتن یک ساختار درختی از آناماتونها پادگیر برای پادگیری هر پارامتر، پارامتر مورد نظر را تعیین کرد. در این ساختار سطوح بالاتر (زدیک به ریشه) یا نکر بازدهی بزرگ و سطوح پایه‌ی، یا نکر بازدهی کوچک هستند و برگها یا نکر مقادیر پارامترها می‌باشند. با توجه به دقت موره انتظار برای هر پارامتر، می‌توان تعداد سطوح ساختار درختی را انتخاب کرد. مقدار هر پارامتر با توجه به مسیر فعلی شده در هر درخت تعیین می‌گردد [۲].

## ۵ - روش Sastry و Thatachar

در این قسمت بشرح روش پیشنهاد شده، تابع توسط  $R$  و  $Thatachar$  را معرفی می‌برازیم. این روش، یک روش براساس آناماتونهای پادگیر با ساختار منفرد است. گروه آناماتونها در یک بازی همکار با پاسخ پکسان با هدف یافتن تابع تمايز بهینه از فضای تابع تمايز موجود شرکت می‌کنند. در این روش تابعی فر ماتریس پاداش بازی در طول پادگیری ذخیره می‌شود. ماتریس پاداش، یک ماتریس چند بعدی است که اعضای آن احتمالهای دسته‌بندی صحیح برای هر مجموعه مقادیر مسکن برای پارامترهای آناماتونهای شرکت کننده در بازی می‌باشند. عمل پادگیری، معنی در همگرا کردن تیم به سمت انتخاب مجموعه احتمال بهینه (تابع تمايز بهینه) می‌نماید. اگر فرض کنیم در مرحله  $k$  ام، عمل  $\alpha_{\frac{n}{n}}$

به از  $= 0.1$  دست متوسط به دست آمده،  $0.5/0.5$  با تعداد متوسط مراحل:

$= 1.08284$  است.

```
for i := 1 to 100 do
  GenerateSamples;
  Repeat  SelectActions;
    DiscResult := Discriminate;
    UpdateGame(DiscResult);
    Until EndOfGame;
  End for
  SaveResults;
end
```

الگوریتم ۱

مسئله نمونه با استفاده از سه الگوریتم بهمنگام سازی فوق بیدامزاری شد. بر  
ساس نتایج بدست آمده، نکات زیر قابل توجه است:

- ۱- الگوریتم  $L_{R\_P}$  قادر به همگرا شدن نمی‌باشد.
- ۲- الگوریتم  $L_{R\_I}$  معموله همگرا شده و قادر به تعیین مقادیر مناسب برای  
بارامترهای توابع تابعی می‌باشد.

۳- همگرا بودن الگوریتم  $L_{R\_EP}$  بستگی به نسبت  $\frac{a}{b}$  دارد. این نسبت باید از پک  
مقدار مشخص  $\gamma$  که به مسئله مورد نظر و محدوده بارامترهای  $a$  و  $b$  بستگی  
دارد، بزرگتر باشد.

۴- از بین دو الگوریتم  $L_{R\_EP}$  و  $L_{R\_I}$ ، الگوریتم  $L_{R\_EP}$  دقت بیشتری تولید  
می‌کند اما بسته به الگوریتم  $L_{R\_I}$ ، دارای سرعت همگرا بودن تری می‌باشد.

در آزمایش تجام شده مقادیر

$$\begin{aligned} & \dots, 0.004, 0.007, 0.005, 0.003, 0.002, 0.001, \\ & \text{برای } a \text{ و مقادیر } b = 0.005, 0.004, 0.003, 0.002, 0.001, \\ & \text{زوجهای } (a, b) \text{ با شرط } 10 > \frac{a}{b}, \text{ الگوریتم اجرا گردید. نتایج بدست آمده نشان}\end{aligned}$$

می‌دهد که برای الگوریتم  $L_{R\_EP}$ ، با کوچکتر شدن  $a$  و  $b$  دقت افزایش می‌باشد  
اما تعداد مراحل مورد نیاز می‌تواند بطور چشمگیری افزایش باید تا آنچه که  
سیستم عملاً نایابدار شده و بجای همگرا شدن به یک نقطه بینه، بین دو یا جند  
نقطه بینه محل نوسان کند. در آزمایش نام مذکور شده، در حالت  
 $a = 0.007, b = 0.005$  سیستم به نایابداری رسید (بعد از  $20,000,000$  مرحله، سیستم همچنان در حال نوسان بود). به همین علت برای  $a < 0.007$   
شرط  $10 > \frac{a}{b}$  روی زوجهای  $(a, b)$  درنظر گرفته شده است.

نمودارهای (۲ - a) و (۲ - b) دقت مستنبتدی و تعداد مراحل مورد نیاز  
جهت همگرا بودن را برای الگوریتمهای فوق به ازاء مقادیر مختلف  $a$  و  $b$  با شرط  
گفته شده را نشان می‌دهند. با توجه به نمودارهای (۲-الف) و (۲-ب) ملاحظه  
می‌شود که با کاهش مقادیر  $a$  و  $b$ ، دقت افزایش می‌باشد ولی در عوض تعداد  
مراحل مورد نیاز بیشتر می‌شود. در نمودارهای فوق به ازاء  $a = b = 0.005$   
دقت و تعداد مراحل برای الگوریتم  $L_{R\_I}$  بسته می‌باشد. برای الگوریتم  
 $L_{R\_I}$ ، به ازاء  $a = 0.002, b = 0.001$  دقت مستنبتدی  $97/416$  با تعداد متوسط مراحل  
 $46641$  و برای الگوریتم  $L_{R\_EP}$  با ازاء  $a = 0.002, b = 0.001$  دقت  $97/750$  با تعداد مراحل متوسط  $60012$  حاصل شده است.

#### ۷- آتماتونهای یادگیری با ساختار ثابت

در این قسمت به بررسی کارایی آتماتونهای یادگیری با ساختار ثابت استناد شده  
در یادگیری توابع تابعی در حل مسئله نمونه می‌پردازیم. آتماتونهای بررسی شده  
در این قسمت عبارتند از: TSetline  $G_{KN,K}$ ، TSetline  $L_{KN,K}$ ، Krinsky و Krylov.  
آناتامتون عمل توسط حالت آناتامتون مشخص می‌شود به این ترتیب که ابتدا  
انتخاب عمل از طرف اولیه  $\Phi(0)$  شروع و عمل  $\alpha(0)$  انتخاب تمحاب می‌کند.  
با دریافت پاسخ عمل از طرف محیط و با توجه به ساختار آناتامتون و نحوه تغییر  
حالت آن، آناتامتون به حالت  $\Phi(1)$  می‌رود و این چرخه ادامه پیدا می‌کند.  
شکل‌های (۲ - الف) تا (۲ - د) گراف تغییر حالت آتماتونهای فوق الذکر را نشان  
می‌دهند.

مدلهای ارائه شده بر ساس مدل محیط  $P$  می‌باشد. در گرافهای مذکور برای  
آناتامتون، در عمل  $\alpha_1, \alpha_2, \alpha_3$  درنظر گرفته شده است. هر ملکین  $N$  حالت دارد  
که عین حافظه آناتامتون نمایه می‌شود. اگر  $\Phi(k)$ ، حالت ملکین در مرحله  
کام، بین  $\Phi_N$  تا  $\Phi_{N+1}$  باشد، عمل  $\alpha_i$  و اگر  $\Phi(k)$  بین  $\Phi_{N+1}$  تا  $\Phi_{N+2}$  باشد،

```

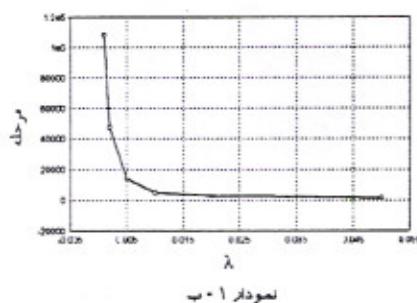
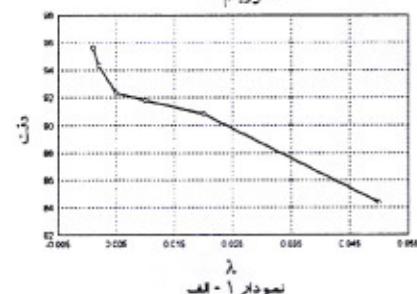
g := random; s := 0; i := 1;
while (i <= r) and not ((g >= s) and (g <= s + p[i])) do s := s+p[i];
SelectedAction := i;

```

الگوریتم ۲

Function EndOfGame : Boolean;  
Result:=True if for All  $A^i$  if  $p[OptimalAction] \geq 0.99$ ;  
End EndOfGame;

الگوریتم ۲



#### ۶- آتماتونهای یادگیری با ساختار متغیر

در این بخش به بررسی کارایی آتماتونهای یادگیری با ساختار متغیر که دارای  
شیوه بهمنگام سازی خطوط احتمالها می‌باشد، می‌پردازیم. اگر در مرحله  $k$  ام،  
عمل  $\alpha_i(k)$  انتخاب شده باشد، بردار احتمال عمل بصورت زیر بهمنگام می‌شود:

الف - با دریافت پاسخ مطلوب :

$$\begin{aligned} P_i(k+1) &= P_i(k) + a[1 - P_i(k)] \\ P_j(k+1) &= (1 - a)P_j(k), \quad j \neq i \end{aligned} \quad (4)$$

ب - با دریافت پاسخ نامطلوب :

$$\begin{aligned} P_i(k+1) &= (1 - b)P_i(k) \\ P_j(k+1) &= \frac{b}{T-1} + (1 - b)P_j(k), \quad j \neq i \end{aligned} \quad (5)$$

و  $a$  و  $b$  بارامترهای یادگیری هستند [۹]. با توجه به روابط (۴) و (۵) با دریافت پاسخ مطلوب از محیط احتمال عمل  
انتخاب شده افزایش یافته و احتمال سایر عملها کاهش می‌باشد و با دریافت پاسخ  
نامطلوب، احتمال عمل انتخاب شده کاهش یافته و احتمال سایر عملها افزایش  
می‌باشد. باید توجه داشت که بهمنگام سازی به گونه‌ای تجام می‌شود که در هر  
مرحله  $k$  داشته باشیم:

$$\forall i \in [1 \dots r] : P_i(k) \geq 0, \sum_{j=1}^r P_j(k) = 1$$

برحسب اینکه مقادیر  $a$  و  $b$  و رابطه آنها نسبت به هم چگونه باشد سه نوع  
الگوریتم بهمنگام سازی حاصل می‌شود که عبارتند از:  
 $1 > a = b = 0 : L_{R\_P}$   
 $1 > a > b = 0 : L_{R\_EP}$   
 $1 > a > 0, b = 0 : L_{R\_I}$

در آنماتونهای با ساختار متغیر از بردار احتمال عمل برای مشخص کردن خاتمه الگوریتم استفاده گردید. در آنماتونهای با ساختار ثابت به دلیل اینکه چنین برداری به طور صریح وجود ندارد، پایه ملاک دیگری برای خاتمه الگوریتم تعریف شود. بدین مفهوم برای هر عمل  $\alpha_i^n$  از آنماتون  $A_i^n$  که دارای  $n$  عمل

است، وزن  $w_i^n$  را با مقدار اولیه  $\frac{1}{T_n}$  تعریف می‌کنیم. شرط خاتمه بازی، رسیدن

وزن عمل بهینه به حد معین خواهد بود  $(w_{i_n}^n \geq w_0)$ . این وزنها در پایان هر مرحله توسط روابط  $W$  به عنوان می‌شوند. در روابط  $W$  فرض بر این است که در مرحله  $K$  اعم عمل  $\alpha_i^n$  انتخاب شده باشد.

$$P_i^{n+1} = \frac{P_i^n (W - 1) + 1}{W} \quad (6)$$

$$P_j^{n+1} = \frac{P_j^n (W - 1)}{W}, \quad j \neq i$$

عددی صحیح است و پارامتر پادگیری نامیده می‌شود. اگر عمل انتخاب شده توسط آنماتون  $A_i^n$  در مرحله  $K$  اعم عمل  $\alpha_i^n$  باشد، با توجه به رابطه (6) پسادگی می‌توان نشان داد:

$$\Delta w_{i_n}^n(k) = w_{i_n}^n(k+1) - w_{i_n}^n(k) = \frac{1 - w_{i_n}^n(k)}{W} \quad (7)$$

$$\sum_{j=1}^n w_{i_n}^n(k) = 1 \quad (8)$$

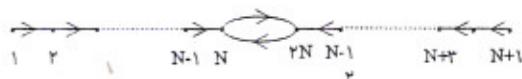
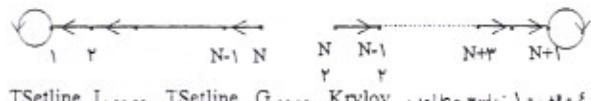
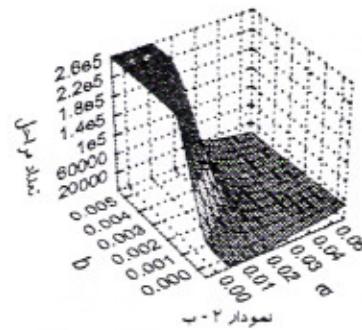
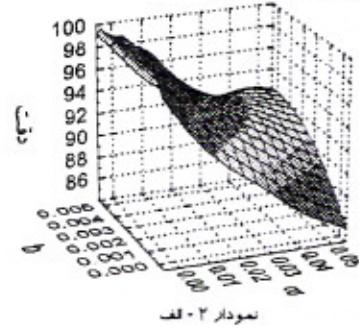
رابطه (7) نشان می‌دهد که با بزرگتر شدن  $W$  سرعت همگرایی کاهش می‌پذد.

پارامتر دیگری که در پادگیری مؤثر است، تعداد نمونهای آموزشی است که در هر مرحله، نتایج تمايز روى آنها انجام می‌شود. وقت تابع بسته آمده در هر مرحله می‌تواند با آزمایش روی  $S$  نمونه از مجموعه نمونهای آموزش سنجیده شود. این دقت به عنوان پاسخ ارسالی محیط به تیم آنماتونها و احتمال مناسب بودن عملهای انتخاب شده در هر مرحله در نظر گرفته می‌شود. اگر فرض شود در هر مرحله  $S$  نمونه دسته‌بندی شوند و تعداد  $S$  نمونه صحیح و

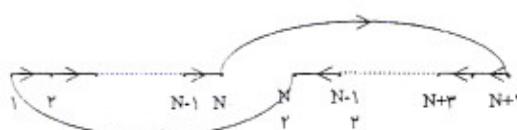
$S_p$  نمونه خلط دسته‌بندی شده باشند، مقدار  $T_p = \frac{S_p}{S}$  به عنوان پاسخ محیط در نظر گرفته می‌شود. به این ترتیب از دید محیط، احتمال درست بودن  $P = 1 - T_p = 1 - \frac{S_p}{S}$  احتمال خلط بودن مجموعه عملهای انتخاب شده می‌باشد. با توجه به اینکه  $S$  مقادیر صحیح بین  $[S, S]$  را اختیار می‌کند، مجموعه خروجی محیط (مجموعه مقادیر ممکن برای  $T_p$ )  $\left\{ \frac{S-1}{S}, \dots, \frac{1}{S} \right\}$  خواهد بود. اگر  $S = 1$  باشد، مجموعه خروجی به  $\{0, 1\}$  تبدیل می‌شود که بیانگر مدل  $P$  برای محیط است و اگر  $S > 1$  باشد، مجموعه خروجی شامل  $+1$  مقدار گستره خواهد بود که بیانگر مدل  $Q$  برای محیط است. در حالت مدل  $Q$ ، مقدار بزرگتر برای پاسخ محیط بیانگر مطابق خواهد بود.

در ادامه این بخش به بررسی پارامترهای مختلف این آنماتونها در حل مسئله نمونه می‌پردازیم. ابتدا تأثیر مقادیر مختلف  $S$  را بررسی می‌کنیم و سپس به بررسی تأثیر عمق خلفه و پارامتر  $W$  می‌پردازیم. شرط پایان الگوریتم در بیان صاری کلیه الگوریتمهای آنماتونهای پادگیر با ساختار ثابت، رسیدن وزن عمل بهینه به بیش از  $99\%$  در نظر گرفته شده است.

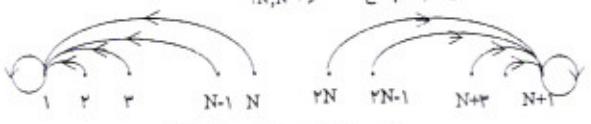
عمل  $\alpha_i^n$  انتخاب می‌شود. عمق خلفه پکی از پارامترهایی است که در کارایی آنماتونهای پادگیر با ساختار ثابت بسیار موثر است. با توجه به گرافهای فوق در صورتیکه عمق حافظه ۱ باشد ( $N = 1$ ) کلیه آنماتونهای فوق (به غیر از مدل Krylov) فرم یکسان، خواهند داشت.



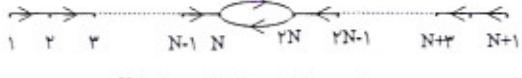
E - 2: پاسخ نامطلوب



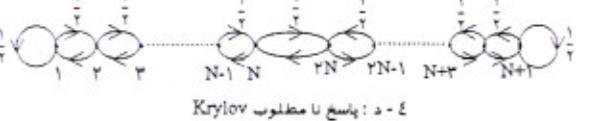
E - 3: پاسخ نامطلوب



E - 4: پاسخ نامطلوب

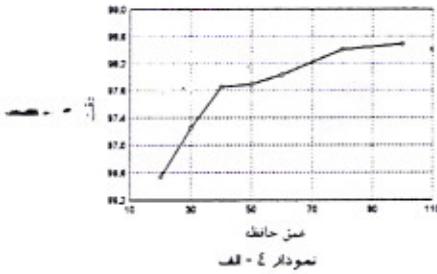


E - 5: پاسخ نامطلوب

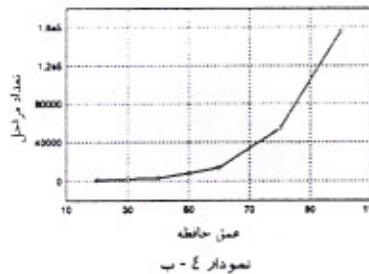


E - 6: پاسخ نامطلوب

حافظه ۱ و W های مختلف کاربری را تحقیق می‌کنیم. نمودارهای (۴-الف) و (۴-ب) نمودارهای دقت مستبندی و تعداد مراحل مورد نیاز به ازای مقادیر مختلف حافظه ۱ و W های مختلف را تحقیق می‌کنیم. نمودارهای (۴-الف) و (۴-ب) نمودارهای دقت مستبندی و تعداد مراحل مورد نیاز به ازای مقادیر مختلف حافظه ۱ و W های مختلف را تحقیق می‌کنند.



نمودار ۴ - الف



نمودار ۴ - ب

همانطور که انتظار می‌رفت با افزایش W دقت افزایش می‌یابد و سرعت همگرایی کاهش می‌یابد. ملاحظه می‌شود که به ازای  $W = 100$  دقت  $W = 98/486$  با تعداد مراحل  $100570$  و به ازای  $W = 80$  دقت  $W = 98/409$  با تعداد مراحل  $52011$  و به ازای  $W = 60$  دقت  $W = 98/377$  با تعداد مراحل  $14001$  بدست می‌آید. به ازای  $W > 60$  سرعت همگرایی به سرعت کم می‌شود (تعداد مراحل مورد نیاز به سرعت افزایش می‌یابد) و دقت به کندی افزایش می‌یابد.

جهت بررسی تأثیر عمق حافظه (N) و W در کاربری مقادیر N و مقادیر W در کاربری (N, W) گوریتم (N) و W در کاربری مقادیر

$N = 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 15$  و  $W = 100, 200, 300, 400, 500, 600, 800, 1000, 1200, 1500$

نظر گرته و به ازای روش‌های مختلف (N, W) گوریتم را اجرا نموده‌ایم.

نمودارهای (۵-الف) تا (۵-ح) نتایج این بررسی را نشان می‌دهند. نمودارهای (۵-الف) و (۵-ب)، نمودارهای دقت و سرعت همگرایی مربوط به آناتامون Tsetline L<sub>KN,K</sub> نمودارهای (۵-ج) و (۵-د) مربوط به آناتامون

Tsetline G<sub>KN,K</sub> ، نمودارهای (۵-۰) و (۵-۰) مربوط به آناتامون Krylov

و نمودارهای (۵-ز) و (۵-ح) مربوط به آناتامون Krinsky

می‌باشد. مشاهده می‌شود که آناتامونها دارای رفتاری تقریباً پیکان می‌باشد. با

افزایش W، همانطور که انتظار می‌رفت دقت افزایش و سرعت همگرایی کاهش می‌یابد و با افزایش عمق حافظه، دقت مستبندی یاپین می‌آید. به ازای

$W = 2$  و  $N = 150$  بهترین دقتها بدست آمده که برای آناتامونها مختلف

عبارتند از: آناتامون Tsetline L<sub>KN,K</sub> دقت  $97/501$  با متوسط تعداد

مراحل  $28526$  ، آناتامون Tsetline G<sub>KN,K</sub> دقت  $97/529$  با تعداد

مراحل  $22146$  ، آناتامون Krinsky دقت  $97/580$  با تعداد مراحل  $26633$  ،

آناتامون Krylov دقت  $97/99$  با تعداد مراحل  $22222$  . با مقایسه این

نمودارها و نتایج بدست آمده برای آناتامونها با عمق حافظه ۱ ، من توان ادعا کرد که آناتامونها با عمق حافظه ۱ ، هم از نظر سرعت و هم از نظر دقت، دارای کاربری بالاتری هستند.

با توجه تغییراتی در آناتامون با عمق حافظه ۱ میتوان دقت و سرعت

همگرایی را بهبود بخشد که چند نمونه از این تغییرات در دو بخش بعدی پیشنهاد شده است.

## ۱ - ۷ - تأثیر حجم نمونه آزمایشی

بر این بخش به بررسی تأثیر پارامتر S بر روی آناتامون L<sub>KN,K</sub> با عمق حافظه ۲ و  $W = 200$  می‌پردازیم. لسلت اصلی برنامه در گوریتم ۱ آمده است. گوریتم ۵ نحوه تغییر حالت مانش را حفظ می‌کند که شامل عمل

انتخاب شده و عمق حافظه عمل می‌باشد و  $\pi$  پاسخ محیط است. در صورت تبیه شدن و رسیدن عمق عمل به ۱ ، عمل دیگری به غیر از عمل فعلی انتخاب می‌شود که این انتخاب می‌تواند بصورت تعادلی انجام گیرد [۹].

گوریتم فوق به ازای مقادیر مختلف گوریتم (۵-الف) و (۵-ب) آمده است. شده و فقط مستبندی و تعداد مراحل مورد نیاز برای همگرایی در نمودارهای (۵-الف) و (۵-ب) آمده است.

Procedure UpdateGame ( $r_e : \text{Real}$ )

Reward := Random  $\leq r_e$  ; Penalty := not Reward;

if Reward then

if State.Depth  $\leq$  MemoryDepth then inc (State.Depth);

End if

if Penalty then

if State.Depth  $> 1$  then dec (State.Depth)

else State.Action := AnotherAction

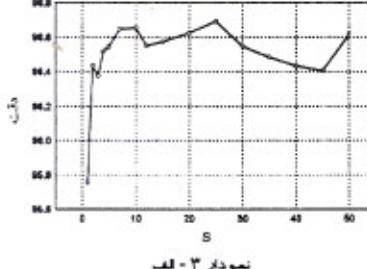
End if

End UpdateGame

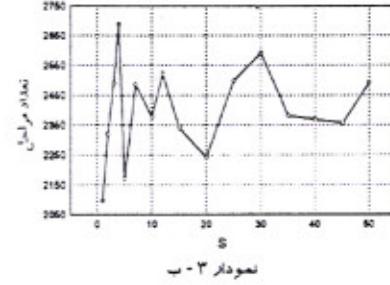
## گوریتم ۵

ملاحظه می‌شود که به ازای  $S = 2$  بهبود زیادی (حدود ۶۱%) در دقت ایجاد شده در حالی که تعداد مراحل حدود ۳۵۰ مرحله افزایش پائمه است. بهترین نتیجه به ازای  $S = 25$  بدست آمده که در آن دقت مستبندی صحیح  $96/690$  با تعداد مراحل ۲۴۹۷ بوده است.

باید توجه داشت که تعداد مراحل الزاماً به معنی تعداد مقایسه‌ها نیست بلکه اگر تعداد مراحل مورد نیاز در هر آزمایش k باشد و تعداد نمونه‌ای مورد آزمایش هر مرحله S باشد، تعداد مقایسه‌ها S.k خواهد بود که در تحلیل سرعت عمل آناتامونها نقش اصلی را بازی می‌کند.



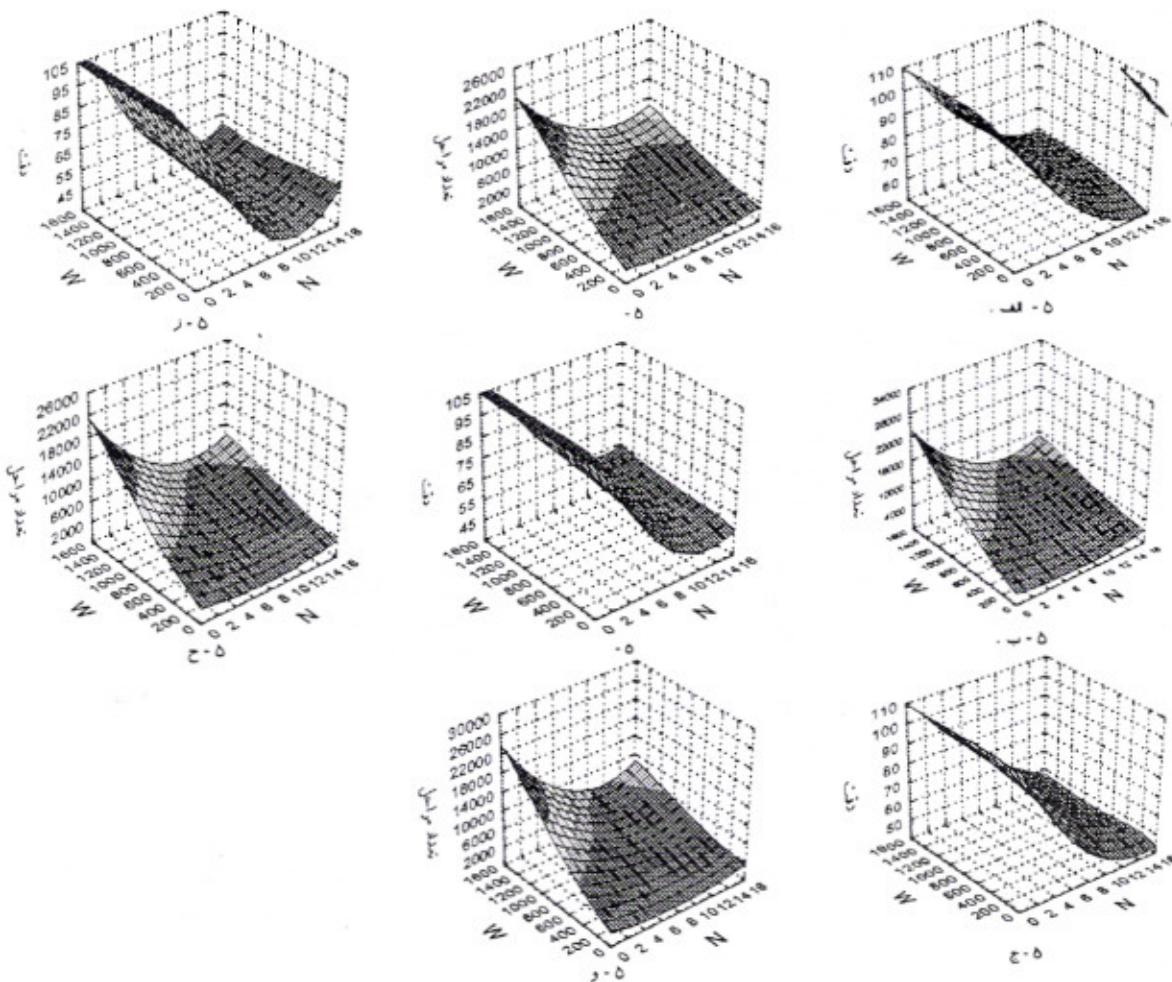
نمودار ۵ - الف



نمودار ۵ - ب

## ۲ - ۷ - تأثیر عمق و پارامتر W

برای عمق حافظه ۱ رفتار بیشتر آناتامونها فوک پیکان می‌باشد و در ضمن محدوده پارامتر W برای عمق ۱ با سایر عمقها متفاوت است، پس اینجا به ازای عمق



پلاش (q<sub>e</sub>) در بخش ۷ تعریف شده‌اند)، احتمالهای تغییر حالت آناتامون (پاداش و تنبیه) را به صورت زیر تعریف می‌کنیم:

- احتمال دریافت پاداش

$$P = \begin{cases} r_e & \text{if } w_{i_n}^n(k) > w_{i_n, \max}^n(k) \\ \frac{w_{i_n}^n(k)}{w_{i_n, \max}^n(k)} & \text{otherwise} \end{cases} \quad (8)$$

- احتمال تنبیه

$$q_e = \begin{cases} q_e & \text{if } w_{i_n}^n(k) < w_{i_n, \max}^n(k) \\ \frac{w_{i_n, \max}^n(k)}{w_{i_n}^n(k)} & \text{otherwise} \end{cases} \quad (9)$$

به طوریکه:

$$w_{i_n, \max}^n(k) = \text{Max} \left\{ w_j^n(k) \mid j = 1, \dots, r_n, j = i_n \right\} \quad (10)$$

در روابط فوق  $w_j^n$  وزن عمل  $j$  است که با استفاده از روابط (۶) بعنوان

#### ۸- آناتامونهای پیشنهادی گروه اول

در این قسمت آناتامونی پیشنهادی می‌دهیم که از دقت و سرعت همگرایی بهتری در حل مسئله شناسایی تابع تعابز برخوردار است. نتایج بدست آمده در بخش قبل نشان داد که عمق حافظه ۱ می‌تواند بهترین پاسخ را در حل مسئله فراهم کند. این نکته بدین معنی است که اجاد سهولت در تغییر عمل می‌تواند باعث هدایت آناتامون به سمت انتخاب عمل بشهود شود. اما اگر بدست آوردن دقت بالا مدنظر باشد، این ایجاد سهولت در تغییر عمل می‌تواند باعث نوسانات اضافی آناتامون شود. در عملهای مختلف و در نتیجه کند شدن سرعت همگرایی آناتامون بین آناتامونهای با عمق حافظه ۱، تغییر عمل با دریافت پاسخ نامطلوب از جانب محیط انجام می‌گیرد یعنی پاداش و تنبیه با توجه به پاسخ محیط تعین می‌شود. با تعابز قائل شدن بین پاسخ محیط و تنبیه و تشریق آناتامون، می‌توان به نتایج بهتری رسید. این تعابز از طریق قرار دادن فیلترهای مناسب بر روی پاسخ محیط بمنظور ضعیفتر و قویتر کردن آن حاصل می‌شود.

فرض کنیم آناتامون  $A^n$  که دارای  $r_n$  عمل است و در مرحله  $k$  عمل  $\alpha_{i_n}^n(k)$  را انتخاب و به محیط اعلام کند. با فرض اینکه  $r_e$  احتمال مناسب بودن عمل  $\alpha_{i_n}^n$  است که با استفاده از روابط (۶) بعنوان

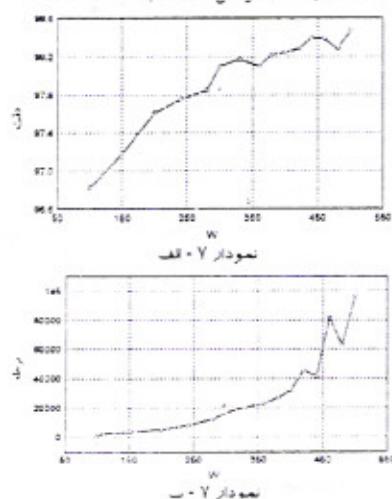
$$p = \begin{cases} r_e & \text{if } PI \geq PMax \\ r_e \times \frac{PI}{PMax} & \text{otherwise} \end{cases} \quad (11)$$

$$q = \begin{cases} q_e & \text{if } PI < PMax \\ q_e \times \frac{PMax}{PI} & \text{otherwise} \end{cases} \quad (12)$$

بطوریکه

$$PI = \sqrt[N]{\prod_{n=1}^N w_{i_n}^{i_n}(k)}, \quad PMax = \sqrt[N]{\prod_{n=1}^N w_{i_n, \max}^{i_n}(k)} \quad (13)$$

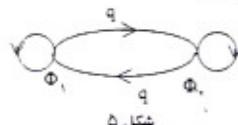
طبق رابطه (۱۳) ،  $PI$  میانگین هندسی وزنهای عملهای انتخاب شده توسط آناتامتون شرکت کنده، در بازی است و  $PMax$  میانگین هندسی حداقل وزن عملهای انتخاب شده است. آناتامتون که طبق روابط (۱۱) و (۱۲) تبیه و تشویق منشود را آناتامتون نوع ۲ نامیم. نمودارهای (۷-الف) و (۷-ب) نتایج بدست آمده از پیداگزی آناتامتون نوع ۲ به ازای مقادیر مختلف  $W$  می‌باشد. با انجام تغییر فوق خطی بردن تعداد مرحله‌ای اینی رفتہ و رفتار آناتامتون آشنازتر و همگرای مجموعه کنترل شده است بطوریکه حداقل دقت  $W = 480/281$  با تعداد مرحله‌ای  $W = 480, 200, 220, 280, 300, 320, 380, 400, 420, 440, 460, 480, 500$  برای آناتامتون آشنازتر و تبیه دیگر را نیز در این راستا آزمایش می‌کنیم و آناتامتون مربوطه را آناتامتونهای نوع ۳ و نوع ۴ می‌نامیم. در آناتامتون نوع ۲، برای تبیه و تشویق مشابه روابط (۱۱) و (۱۲) عمل می‌کنیم با این تفاوت که به جای  $PI$  از  $W_{i_n}$  یعنی وزن عمل انتخاب شده استفاده گردیده است.



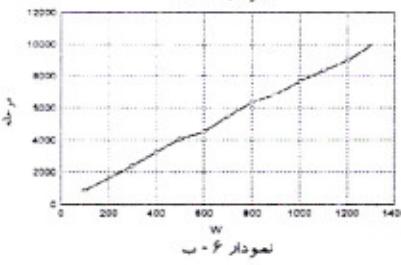
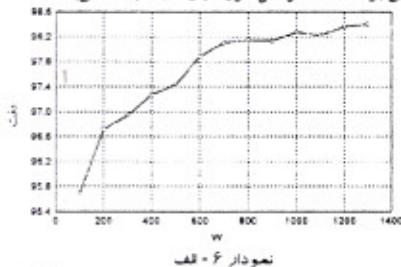
در شیوه تبیه و تشویق دیگر را نیز در این راستا آزمایش می‌کنیم و آناتامتونهای مربوطه را آناتامتونهای نوع ۳ و نوع ۴ می‌نامیم. در آناتامتون نوع ۲، برای تبیه و تشویق مشابه روابط (۱۱) و (۱۲) عمل می‌کنیم با این تفاوت که به جای  $PI$  از  $W_{i_n}$  یعنی وزن عمل انتخاب شده استفاده گردیده است. نتایج این روش در نمودارهای (۷-الف) و (۷-ب) آمده است. شیوه دیگر، استفاده از میانگین حسابی به جای میانگین هندسی و استفاده از  $W_{i_n}$  در روابط (۱۱) و (۱۲) به جای  $PI$  می‌باشد. به این ترتیب  $PMax$  با استفاده از رابطه (۱۲) محاسبه می‌شود. نتایج روش دوم در نمودارهای (۷-ج) و (۷-د) آمده است. آناتامتون نوع ۳ به ازای مقادیر آناتامتون  $W = 480, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300$  و آناتامتونهای نوع ۴ به ازای مقادیر آناتامتون  $W = 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400$  آزمایش شده‌اند. حداقل دقت بدست آمده در روش اول به ازای  $W = 700$  برابر  $98/422$  با تعداد مرحله‌ای ۸۸۳۳ و در روش دوم به ازای  $W = 550$ .

منشود و  $\frac{W}{PMax}$  حداقل وزن عملهای انتخاب شده می‌باشد. طبق رابطه

(۸) اگر آناتامتون در حالتی باشد که وزن عمل انتخاب شده در آن حالت، کمتر و حداقل وزن سایر عملها باشد و پاسخ مطلوب دریافت کرده باشد، با احتمال کمتری نسبت به پاسخ محیط تشویق می‌شود. این موضوع بدین معنی است که احتملاً عمل انتخاب شده فعلی، عمل بهینه نمی‌باشد و بنابراین پاسخ محیط تضعیف می‌شود. طبق رابطه (۹) اگر آناتامتون در حالتی باشد که وزن عمل انتخاب شده در آن حالت، پیشترین وزن باشد و پاسخ نامطلوب دریافت کرده باشد، احتمال تبیه شدن آن کمتر از پاسخ محیط در نظر گرفته می‌شود به این معنی که احتملاً عمل انتخاب شده، عمل بهینه بوده اما در مرحله‌ای باعث برداشت نامطلوب از جانب محیط شده است، بنابراین احتمال تبیه آناتامتون تضییف می‌شود، به این ترتیب روابط (۸) تا (۱۰) با ایجاد تبایز بین پاسخ محیط و تبیه و تشویق آناتامتون، می‌توانند از نوسانات اضافی آناتامتون جلوگیری کنند. آناتامتون که طبق روابط (۸) تا (۱۰) تغییر حالت می‌دهد را آناتامتون نوع ۱ می‌نامیم. در آزمایش تجام شده از آناتامتونهای با عمق حافظه ۱ و پاداش و تبیه طبق روابط (۸) تا (۱۰) مستفاده گردیده است. شکل ۵ گراف تغییر حالت آناتامتون مورود نظر را برای دو عمل نشان می‌دهد.



نمودارهای (۶-الف) و (۶-ب) نتایج دقت و تعداد مرحله‌ای مورد نیاز را برای آناتامتون نوع ۱ به ازای مقادیر مختلف  $W$  نشان می‌دهند. مشاهده می‌شود که نتایج بدست آمده از سرعت و دقت بهتری برخوردار می‌باشند. به ازای  $W = 700$  دقت مستحبندی  $98/102$  با متوسط تعداد مرحله‌ای  $5203$  و به ازای  $W = 1300$  دقت  $W = 98/395$  با متوسط  $9899$  بدست آمده است. نکته جالب دیگری که در نمودار (۶-ب) به چشم می‌خورد خطی بردن تعداد مرحله‌ای مورد نیاز نسبت به  $W$  می‌باشد.



در قدم بعدی سعی در ایجاد ارتباط محکمترین آناتامتونهای شرکت کننده در بازی می‌کنیم به این امید که نتایج بهتری بدست آید. اگر فرض کنیم  $N$  آناتامتون  $A^1, A^2, \dots, A^N$  در یک بازی با پاسخ پکان شرکت کرده‌اند و عمل انتخاب شده در مرحله  $k$  توسط آناتامتون  $A^B$ ،  $B \in \{1, 2, \dots, N\}$  انجام شده باشد احتمالهای پاداش و تبیه آناتامتون را به صورت زیر تعریف می‌کنیم :

پاسخ مناسب  $\tau_e$  و پاسخ نامناسب  $q_e$  از جانب محیط در شکل (۶ - a) و (۶ - b) نشان داده شده است.

در گرافهای فوق فرض بر این است که در مرحله‌ای، آتاماتون در حالت  $\Phi_{\alpha_1}$  قرار دارد، به این ترتیب در این مرحله، عمل  $\alpha_1$  انتخاب می‌شود. با دریافت پاسخ مثبت  $\tau_e$  از جانب محیط آتاماتون با احتمال  $p$  تشویق می‌شود (که با  $\tau_e$  ارتباط دارد) که در این صورت به عمق عمل  $\alpha_1$ ، یک واحد اضافه می‌شود (آتاماتون به حالت  $\Phi_{\alpha_1 + 1}$  می‌رود). در صورتیکه آتاماتون در حالت  $\Phi_N$  باشد در همان حالت باقی می‌ماند. در صورت انجام عمل تشویق، فرض می‌کنیم حالت قبلی آتاماتون در انتخاب عمل  $\alpha_1$ ،  $\alpha_2$  باشد، در این صورت با احتمال  $q$  تبیه می‌شود که در این صورت، در صورتیکه عمق قبلی  $\Phi_{N+1}$  باشد، یک واحد از عمق آن کم می‌شود.

با دریافت پاسخ منفی  $q_e$  از جانب محیط ( $q_e = 1 - \tau_e$ ) آتاماتون با احتمال  $q'$  تبیه می‌شود که در این صورت، حالت فعلی آن به عنوان حالت قبلی عمل  $\alpha_1$  ذخیره شده و آتاماتون به حالت قبلی عمل  $\alpha_1$  (که در شکل ۶ - c) است) تغییر حالت می‌دهد. عمق اولیه برای کلیه عملها در شروع کار آتاماتون باشد مقدار دهن اولیه شود. این مقدار دهن اولیه می‌تواند بصورت  $\frac{N}{T}$  باشد که  $N$  عمق حافظه آتاماتون و  $T$  تعداد عملهای آن است.

برای ارتباط  $\tau_e$ ,  $q_e$ ,  $p$ ,  $q$ ,  $q'$  با  $p$ ,  $q$ ,  $p$ ,  $q$ ,  $q'$  میتوان شیوه‌های مختلفی در نظر گرفت. بطور مثال اگر فرض کنیم آتاماتون  $A_n^B$  با عمق حافظه  $r_n$  دارای  $r_n$  عمل باشد و در مرحله  $k$  عمل  $k$  را در عمق  $w_{i_n}^n(k)$  انتخاب کرده باشد،

اختهای تبیه و تشویق را به صورت زیر تعریف می‌کنیم:

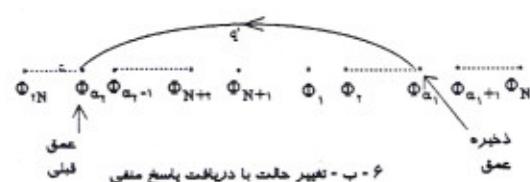
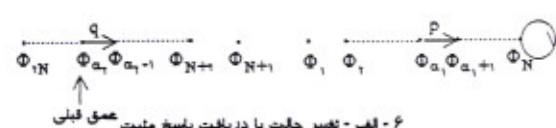
$$p = \begin{cases} r_e & \text{if } w_{i_n}^n \geq w_{i_n, \max}^n(k) \\ r_e \times \frac{w_{i_n}^n}{w_{i_n, \max}^n(k)} & \text{otherwise} \end{cases} \quad (15)$$

$$q = w_j^n(k) \quad j = i_n \quad (16)$$

$$q' = \begin{cases} q_e & \text{if } w_{i_n}^n(k) < w_{i_n, \max}^n(k) \\ q_e \times \frac{w_{i_n, \max}^n(k)}{w_{i_n}^n(k)} & \text{otherwise} \end{cases} \quad (17)$$

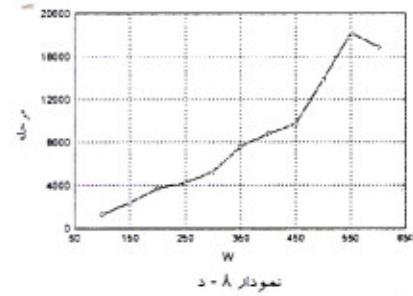
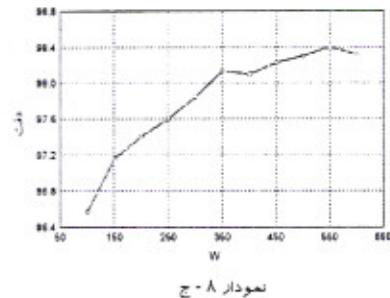
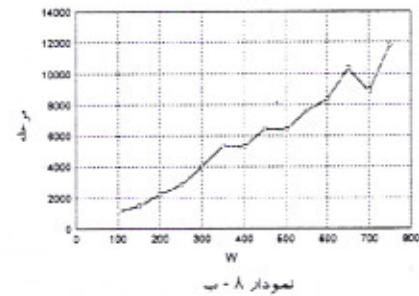
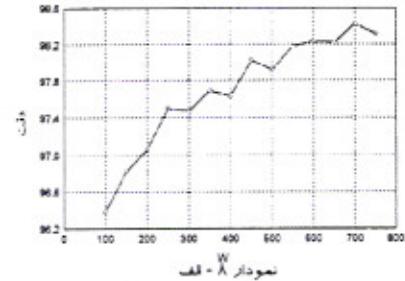
$$w_j^n(k) = \frac{m_j^n(k)}{N}, \quad j = 1, \dots, r_n \quad (18)$$

$$w_{i_n, \max}^n(k) = \text{Max} \left\{ w_j^n(k) \mid j = 1 \dots r_n, j = i_n \right\} \quad (19)$$



۶- آبیمن با تعداد مراحل ۱۸۱۵۷ می‌باشد.

$$\text{PMAX} = \frac{\sum_{k=1}^N w_{i_n, \max}^n(k)}{N} \quad (12)$$



## ۹ - آتاماتونهای پیشنهادی گروه دوم

در این قسم آتاماتون پیشنهاد می‌دهیم که وزنهای  $w_j^n$  در ساختار آتاماتون گنجانده شده است. آتاماتون پیشنهادی دارای دو عمل است که برای جند عمل نیز برخشن قابل تعمیم است.

آتاماتونی با در عمل  $\alpha_1, \alpha_2, \dots, \alpha_{2N}$  حافظه  $\Phi_{i,N}, \Phi_{i+1}, \dots, \Phi_{i+N}$  را در نظر بگیرید.  $\Phi_i$  تا  $\Phi_N$  مجموعه حلال انتخاب عمل  $\alpha_1$  و  $\alpha_{N+1}$  تا  $\alpha_{i+N}$  را مجموعه حلال انتخاب عمل  $\alpha_2$  می‌باشد. گراف تغییر حالت آتاماتون با آیند

$$p = \begin{cases} r_e & \text{if } PI = PMax \\ r_e \times \frac{PI}{PMax} & \text{otherwise} \end{cases} \quad (20)$$

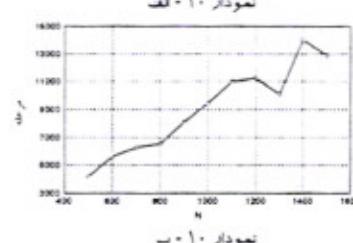
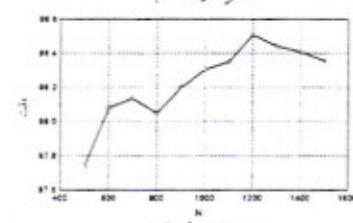
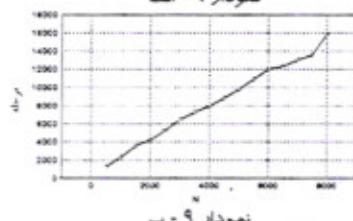
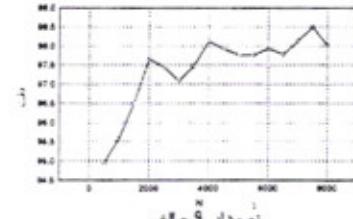
$$q' = \begin{cases} q_e & \text{if } PI < PMax \\ q_e \times \frac{PMax}{PI} & \text{otherwise} \end{cases} \quad (21)$$

بطوریکه:

$$PI = \sqrt[N]{\prod_{n=1}^N w_{i_n}^n(k)}, PMax = \sqrt[N]{\prod_{n=1}^N w_{i_n, \max}^n(k)} \quad (22)$$

احتمال  $q$  طبق ربطه (16) و  $w_{i_n}^n$ ,  $w_{i_n, \max}^n$  طبق ربطه (18) و

(14) محاسبه می‌شوند. در این نوع آناتامون که آناتامون نوع ۶ نامیده می‌شود سعی شده است که ارتباط محکمتری بین آناتامونهای شرکت گشته در بازی برقرار شود. نتایج آزمایش برای این روش برای عمقهای ۵۰۰, ۶۰۰, ۷۰۰, ۸۰۰, ۹۰۰, ۱۰۰۰, ۱۱۰۰, ۱۲۰۰, ۱۳۰۰, ۱۴۰۰, ۱۵۰۰ نمودارهای (۹ - a) و (۹ - b) (آلفا و بتا) می‌باشد. با توجه به نمودارهای فوق مشاهده می‌شود که دقت و سرعت تا حدودی قابل افزایش باقاعد است. حداقل دقت بینست آمدۀ به ازای عمق حافظه ۱۲۰۰ برای ۹۸/۰۰۸ با تعداد مراحل ۱۰۹۹۷ می‌باشد.



### ۱۰ - نتیجه گیری

در بخش‌های قبل روش ارائه شده توسط Thatachar و Sastry در شناسایی نتایج تمايز و همچنین آناتامونهای پادگیر با ساختار متغیر و ثابت استاندارد پیاده سازی شد و کارایی هر یک از روش‌های فوق در شناسایی نتایج

ربطه (18) برای هر عمل وزنی برای نسبت عمق عمل به عمق حافظه آناتامون تعريف می‌کند. ربطه (19) حداکثر وزن سایر عملهای انتخاب نشده برای آناتامون را تعريف می‌کند و ربطه (15) تا (17) احتمالهای تنبیه و تشویق آناتامون را محاسبه می‌کنند. آناتامون که به صورت فوق تغییر حالت می‌دهد را آناتامون نوع ۵ می‌نامیم.

سکلت اصلی برنامه شبیه ساز در الگوریتم ۱ آمدۀ است. در الگوریتم ۶ رویه Initialize که مقدار دهن اولیه عمق را تجاهم می‌دهد، رویه UpdateGame که نحوه تغییر حالت آناتامون را پیاده سازی می‌کند و رویه EndOfGame که پایان بازی را مشخص می‌کند، برای هر یک بازی مشکل از N آناتامون، آورده شده است.

Procedure Initialize;

for  $i := 1$  to  $r_n$  do  $m(i) := N_n / r_n$ 

End Initialize

Procedure UpdateGame ( $r_e$ ) $q_e := 1 - r_e$ ;  $w_i := m(\text{State.Action}) / N_n$ ; $w_{i,\max} := \text{Max}(m(j) / N_n | j = i)$ if  $w_i \geq w_{i,\max}$  then Reward := Random  $\leq r_e$ else Reward := Random  $\leq r_e * w_i / w_{i,\max}$ 

Penalty := not Reward;

if Penalty then

if  $w_i \leq w_{i,\max}$  then Penalty := Random  $\leq q_e$ else Penalty := Random  $\leq q_e * w_{i,\max} / w_i$ 

End if { Penalty }

if Reward then

if State.Depth  $< N_n$  then inc (State.Depth);for  $i := 1$  to  $r_n$  and  $i \neq \text{State.Action}$  doif ( $m(i) > 1$ ) and (Random  $\leq m(i) / N_n$ ) then dec ( $m(i)$ );

End for

End if; { Reward }

if Penalty then

m (State.Action) := State.Depth;

State.Action := Guess Another Action;

State.Depth := m (State.Action);

End if; { Penalty }

End UpdateGame;

Function EndOfGame : Boolean;

Result := True if for All  $n$  in [1..N]  $m(\text{Action.State}_n) / N_n \geq 0.99$ ;

End EndOfGame;

الگوریتم ۶

الگوریتم ۶ برای عمقهای حافظه ( $N_n$ ) مختلف ۵۰۰, ۶۰۰, ۷۰۰, ۸۰۰, ۹۰۰, ۱۰۰۰, ۱۱۰۰, ۱۲۰۰, ۱۳۰۰, ۱۴۰۰, ۱۵۰۰, ۲۰۰۰, ۲۵۰۰, ۳۰۰۰, ۳۵۰۰, ۴۰۰۰, ۴۵۰۰, ۵۰۰۰, ۵۵۰۰, ۶۰۰۰, ۶۵۰۰, ۷۰۰۰, ۷۵۰۰, ۸۰۰۰, ۸۵۰۰, ۹۰۰۰, ۹۵۰۰, ۱۰۰۰۰, ۱۰۵۰۰, ۱۱۰۰۰, ۱۱۵۰۰, ۱۲۰۰۰, ۱۲۵۰۰, ۱۳۰۰۰, ۱۳۵۰۰, ۱۴۰۰۰, ۱۴۵۰۰, ۱۵۰۰۰, ۱۵۵۰۰, ۱۶۰۰۰, ۱۶۵۰۰, ۱۷۰۰۰, ۱۷۵۰۰, ۱۸۰۰۰, ۱۸۵۰۰, ۱۹۰۰۰, ۱۹۵۰۰, ۲۰۰۰۰, ۲۰۵۰۰, ۲۱۰۰۰, ۲۱۵۰۰, ۲۲۰۰۰, ۲۲۵۰۰, ۲۳۰۰۰, ۲۳۵۰۰, ۲۴۰۰۰, ۲۴۵۰۰, ۲۵۰۰۰, ۲۵۵۰۰, ۲۶۰۰۰, ۲۶۵۰۰, ۲۷۰۰۰, ۲۷۵۰۰, ۲۸۰۰۰, ۲۸۵۰۰, ۲۹۰۰۰, ۲۹۵۰۰, ۳۰۰۰۰, ۳۰۵۰۰, ۳۱۰۰۰, ۳۱۵۰۰, ۳۲۰۰۰, ۳۲۵۰۰, ۳۳۰۰۰, ۳۳۵۰۰, ۳۴۰۰۰, ۳۴۵۰۰, ۳۵۰۰۰, ۳۵۵۰۰, ۳۶۰۰۰, ۳۶۵۰۰, ۳۷۰۰۰, ۳۷۵۰۰, ۳۸۰۰۰, ۳۸۵۰۰, ۳۹۰۰۰, ۳۹۵۰۰, ۴۰۰۰۰, ۴۰۵۰۰, ۴۱۰۰۰, ۴۱۵۰۰, ۴۲۰۰۰, ۴۲۵۰۰, ۴۳۰۰۰, ۴۳۵۰۰, ۴۴۰۰۰, ۴۴۵۰۰, ۴۵۰۰۰, ۴۵۵۰۰, ۴۶۰۰۰, ۴۶۵۰۰, ۴۷۰۰۰, ۴۷۵۰۰, ۴۸۰۰۰, ۴۸۵۰۰, ۴۹۰۰۰, ۴۹۵۰۰, ۵۰۰۰۰, ۵۰۵۰۰, ۵۱۰۰۰, ۵۱۵۰۰, ۵۲۰۰۰, ۵۲۵۰۰, ۵۳۰۰۰, ۵۳۵۰۰, ۵۴۰۰۰, ۵۴۵۰۰, ۵۵۰۰۰, ۵۵۵۰۰, ۵۶۰۰۰, ۵۶۵۰۰, ۵۷۰۰۰, ۵۷۵۰۰, ۵۸۰۰۰, ۵۸۵۰۰, ۵۹۰۰۰, ۵۹۵۰۰, ۶۰۰۰۰, ۶۰۵۰۰, ۶۱۰۰۰, ۶۱۵۰۰, ۶۲۰۰۰, ۶۲۵۰۰, ۶۳۰۰۰, ۶۳۵۰۰, ۶۴۰۰۰, ۶۴۵۰۰, ۶۵۰۰۰, ۶۵۵۰۰, ۶۶۰۰۰, ۶۶۵۰۰, ۶۷۰۰۰, ۶۷۵۰۰, ۶۸۰۰۰, ۶۸۵۰۰, ۶۹۰۰۰, ۶۹۵۰۰, ۷۰۰۰۰, ۷۰۵۰۰, ۷۱۰۰۰, ۷۱۵۰۰, ۷۲۰۰۰, ۷۲۵۰۰, ۷۳۰۰۰, ۷۳۵۰۰, ۷۴۰۰۰, ۷۴۵۰۰, ۷۵۰۰۰, ۷۵۵۰۰, ۷۶۰۰۰, ۷۶۵۰۰, ۷۷۰۰۰, ۷۷۵۰۰, ۷۸۰۰۰, ۷۸۵۰۰, ۷۹۰۰۰, ۷۹۵۰۰, ۸۰۰۰۰, ۸۰۵۰۰, ۸۱۰۰۰, ۸۱۵۰۰, ۸۲۰۰۰, ۸۲۵۰۰, ۸۳۰۰۰, ۸۳۵۰۰, ۸۴۰۰۰, ۸۴۵۰۰, ۸۵۰۰۰, ۸۵۵۰۰, ۸۶۰۰۰, ۸۶۵۰۰, ۸۷۰۰۰, ۸۷۵۰۰, ۸۸۰۰۰, ۸۸۵۰۰, ۸۹۰۰۰, ۸۹۵۰۰, ۹۰۰۰۰, ۹۰۵۰۰, ۹۱۰۰۰, ۹۱۵۰۰, ۹۲۰۰۰, ۹۲۵۰۰, ۹۳۰۰۰, ۹۳۵۰۰, ۹۴۰۰۰, ۹۴۵۰۰, ۹۵۰۰۰, ۹۵۵۰۰, ۹۶۰۰۰, ۹۶۵۰۰, ۹۷۰۰۰, ۹۷۵۰۰, ۹۸۰۰۰, ۹۸۵۰۰, ۹۹۰۰۰, ۹۹۵۰۰, ۱۰۰۰۰۰, ۱۰۰۵۰۰, ۱۰۱۰۰۰, ۱۰۱۵۰۰, ۱۰۲۰۰۰, ۱۰۲۵۰۰, ۱۰۳۰۰۰, ۱۰۳۵۰۰, ۱۰۴۰۰۰, ۱۰۴۵۰۰, ۱۰۵۰۰۰, ۱۰۵۵۰۰, ۱۰۶۰۰۰, ۱۰۶۵۰۰, ۱۰۷۰۰۰, ۱۰۷۵۰۰, ۱۰۸۰۰۰, ۱۰۸۵۰۰, ۱۰۹۰۰۰, ۱۰۹۵۰۰, ۱۱۰۰۰۰, ۱۱۰۵۰۰, ۱۱۱۰۰۰, ۱۱۱۵۰۰, ۱۱۲۰۰۰, ۱۱۲۵۰۰, ۱۱۳۰۰۰, ۱۱۳۵۰۰, ۱۱۴۰۰۰, ۱۱۴۵۰۰, ۱۱۵۰۰۰, ۱۱۵۵۰۰, ۱۱۶۰۰۰, ۱۱۶۵۰۰, ۱۱۷۰۰۰, ۱۱۷۵۰۰, ۱۱۸۰۰۰, ۱۱۸۵۰۰, ۱۱۹۰۰۰, ۱۱۹۵۰۰, ۱۲۰۰۰۰, ۱۲۰۵۰۰, ۱۲۱۰۰۰, ۱۲۱۵۰۰, ۱۲۲۰۰۰, ۱۲۲۵۰۰, ۱۲۳۰۰۰, ۱۲۳۵۰۰, ۱۲۴۰۰۰, ۱۲۴۵۰۰, ۱۲۵۰۰۰, ۱۲۵۵۰۰, ۱۲۶۰۰۰, ۱۲۶۵۰۰, ۱۲۷۰۰۰, ۱۲۷۵۰۰, ۱۲۸۰۰۰, ۱۲۸۵۰۰, ۱۲۹۰۰۰, ۱۲۹۵۰۰, ۱۳۰۰۰۰, ۱۳۰۵۰۰, ۱۳۱۰۰۰, ۱۳۱۵۰۰, ۱۳۲۰۰۰, ۱۳۲۵۰۰, ۱۳۳۰۰۰, ۱۳۳۵۰۰, ۱۳۴۰۰۰, ۱۳۴۵۰۰, ۱۳۵۰۰۰, ۱۳۵۵۰۰, ۱۳۶۰۰۰, ۱۳۶۵۰۰, ۱۳۷۰۰۰, ۱۳۷۵۰۰, ۱۳۸۰۰۰, ۱۳۸۵۰۰, ۱۳۹۰۰۰, ۱۳۹۵۰۰, ۱۴۰۰۰۰, ۱۴۰۵۰۰, ۱۴۱۰۰۰, ۱۴۱۵۰۰, ۱۴۲۰۰۰, ۱۴۲۵۰۰, ۱۴۳۰۰۰, ۱۴۳۵۰۰, ۱۴۴۰۰۰, ۱۴۴۵۰۰, ۱۴۵۰۰۰, ۱۴۵۵۰۰, ۱۴۶۰۰۰, ۱۴۶۵۰۰, ۱۴۷۰۰۰, ۱۴۷۵۰۰, ۱۴۸۰۰۰, ۱۴۸۵۰۰, ۱۴۹۰۰۰, ۱۴۹۵۰۰, ۱۵۰۰۰۰, ۱۵۰۵۰۰, ۱۵۱۰۰۰, ۱۵۱۵۰۰, ۱۵۲۰۰۰, ۱۵۲۵۰۰, ۱۵۳۰۰۰, ۱۵۳۵۰۰, ۱۵۴۰۰۰, ۱۵۴۵۰۰, ۱۵۵۰۰۰, ۱۵۵۵۰۰, ۱۵۶۰۰۰, ۱۵۶۵۰۰, ۱۵۷۰۰۰, ۱۵۷۵۰۰, ۱۵۸۰۰۰, ۱۵۸۵۰۰, ۱۵۹۰۰۰, ۱۵۹۵۰۰, ۱۶۰۰۰۰, ۱۶۰۵۰۰, ۱۶۱۰۰۰, ۱۶۱۵۰۰, ۱۶۲۰۰۰, ۱۶۲۵۰۰, ۱۶۳۰۰۰, ۱۶۳۵۰۰, ۱۶۴۰۰۰, ۱۶۴۵۰۰, ۱۶۵۰۰۰, ۱۶۵۵۰۰, ۱۶۶۰۰۰, ۱۶۶۵۰۰, ۱۶۷۰۰۰, ۱۶۷۵۰۰, ۱۶۸۰۰۰, ۱۶۸۵۰۰, ۱۶۹۰۰۰, ۱۶۹۵۰۰, ۱۷۰۰۰۰, ۱۷۰۵۰۰, ۱۷۱۰۰۰, ۱۷۱۵۰۰, ۱۷۲۰۰۰, ۱۷۲۵۰۰, ۱۷۳۰۰۰, ۱۷۳۵۰۰, ۱۷۴۰۰۰, ۱۷۴۵۰۰, ۱۷۵۰۰۰, ۱۷۵۵۰۰, ۱۷۶۰۰۰, ۱۷۶۵۰۰, ۱۷۷۰۰۰, ۱۷۷۵۰۰, ۱۷۸۰۰۰, ۱۷۸۵۰۰, ۱۷۹۰۰۰, ۱۷۹۵۰۰, ۱۸۰۰۰۰, ۱۸۰۵۰۰, ۱۸۱۰۰۰, ۱۸۱۵۰۰, ۱۸۲۰۰۰, ۱۸۲۵۰۰, ۱۸۳۰۰۰, ۱۸۳۵۰۰, ۱۸۴۰۰۰, ۱۸۴۵۰۰, ۱۸۵۰۰۰, ۱۸۵۵۰۰, ۱۸۶۰۰۰, ۱۸۶۵۰۰, ۱۸۷۰۰۰, ۱۸۷۵۰۰, ۱۸۸۰۰۰, ۱۸۸۵۰۰, ۱۸۹۰۰۰, ۱۸۹۵۰۰, ۱۹۰۰۰۰, ۱۹۰۵۰۰, ۱۹۱۰۰۰, ۱۹۱۵۰۰, ۱۹۲۰۰۰, ۱۹۲۵۰۰, ۱۹۳۰۰۰, ۱۹۳۵۰۰, ۱۹۴۰۰۰, ۱۹۴۵۰۰, ۱۹۵۰۰۰, ۱۹۵۵۰۰, ۱۹۶۰۰۰, ۱۹۶۵۰۰, ۱۹۷۰۰۰, ۱۹۷۵۰۰, ۱۹۸۰۰۰, ۱۹۸۵۰۰, ۱۹۹۰۰۰, ۱۹۹۵۰۰, ۲۰۰۰۰۰, ۲۰۰۵۰۰, ۲۰۱۰۰۰, ۲۰۱۵۰۰, ۲۰۲۰۰۰, ۲۰۲۵۰۰, ۲۰۳۰۰۰, ۲۰۳۵۰۰, ۲۰۴۰۰۰, ۲۰۴۵۰۰, ۲۰۵۰۰۰, ۲۰۵۵۰۰, ۲۰۶۰۰۰, ۲۰۶۵۰۰, ۲۰۷۰۰۰, ۲۰۷۵۰۰, ۲۰۸۰۰۰, ۲۰۸۵۰۰, ۲۰۹۰۰۰, ۲۰۹۵۰۰, ۲۱۰۰۰۰, ۲۱۰۵۰۰, ۲۱۱۰۰۰, ۲۱۱۵۰۰, ۲۱۲۰۰۰, ۲۱۲۵۰۰, ۲۱۳۰۰۰, ۲۱۳۵۰۰, ۲۱۴۰۰۰, ۲۱۴۵۰۰, ۲۱۵۰۰۰, ۲۱۵۵۰۰, ۲۱۶۰۰۰, ۲۱۶۵۰۰, ۲۱۷۰۰۰, ۲۱۷۵۰۰, ۲۱۸۰۰۰, ۲۱۸۵۰۰, ۲۱۹۰۰۰, ۲۱۹۵۰۰, ۲۲۰۰۰۰, ۲۲۰۵۰۰, ۲۲۱۰۰۰, ۲۲۱۵۰۰, ۲۲۲۰۰۰, ۲۲۲۵۰۰, ۲۲۳۰۰۰, ۲۲۳۵۰۰, ۲۲۴۰۰۰, ۲۲۴۵۰۰, ۲۲۵۰۰۰, ۲۲۵۵۰۰, ۲۲۶۰۰۰, ۲۲۶۵۰۰, ۲۲۷۰۰۰, ۲۲۷۵۰۰, ۲۲۸۰۰۰, ۲۲۸۵۰۰, ۲۲۹۰۰۰, ۲۲۹۵۰۰, ۲۳۰۰۰۰, ۲۳۰۵۰۰, ۲۳۱۰۰۰, ۲۳۱۵۰۰, ۲۳۲۰۰۰, ۲۳۲۵۰۰, ۲۳۳۰۰۰, ۲۳۳۵۰۰, ۲۳۴۰۰۰, ۲۳۴۵۰۰, ۲۳۵۰۰۰, ۲۳۵۵۰۰, ۲۳۶۰۰۰, ۲۳۶۵۰۰, ۲۳۷۰۰۰, ۲۳۷۵۰۰, ۲۳۸۰۰۰, ۲۳۸۵۰۰, ۲۳۹۰۰۰, ۲۳۹۵۰۰, ۲۴۰۰۰۰, ۲۴۰۵۰۰, ۲۴۱۰۰۰, ۲۴۱۵۰۰, ۲۴۲۰۰۰, ۲۴۲۵۰۰, ۲۴۳۰۰۰, ۲۴۳۵۰۰, ۲۴۴۰۰۰, ۲۴۴۵۰۰, ۲۴۵۰۰۰, ۲۴۵۵۰۰, ۲۴۶۰۰۰, ۲۴۶۵۰۰, ۲۴۷۰۰۰, ۲۴۷۵۰۰, ۲۴۸۰۰۰, ۲۴۸۵۰۰, ۲۴۹۰۰۰, ۲۴۹۵۰۰, ۲۵۰۰۰۰, ۲۵۰۵۰۰, ۲۵۱۰۰۰, ۲۵۱۵۰۰, ۲۵۲۰۰۰, ۲۵۲۵۰۰, ۲۵۳۰۰۰, ۲۵۳۵۰۰, ۲۵۴۰۰۰, ۲۵۴۵۰۰, ۲۵۵۰۰۰, ۲۵۵۵۰۰, ۲۵۶۰۰۰, ۲۵۶۵۰۰, ۲۵۷۰۰۰, ۲۵۷۵۰۰, ۲۵۸۰۰۰, ۲۵۸۵۰۰, ۲۵۹۰۰۰, ۲۵۹۵۰۰, ۲۶۰۰۰۰, ۲۶۰۵۰۰, ۲۶۱۰۰۰, ۲۶۱۵۰۰, ۲۶۲۰۰۰, ۲۶۲۵۰۰, ۲۶۳۰۰۰, ۲۶۳۵۰۰, ۲۶۴۰۰۰, ۲۶۴۵۰۰, ۲۶۵۰۰۰, ۲۶۵۵۰۰, ۲۶۶۰۰۰, ۲۶۶۵۰۰, ۲۶۷۰۰۰, ۲۶۷۵۰۰, ۲۶۸۰۰۰, ۲۶۸۵۰۰, ۲۶۹۰۰۰, ۲۶۹۵۰۰, ۲۷۰۰۰۰, ۲۷۰۵۰۰, ۲۷۱۰۰۰, ۲۷۱۵۰۰, ۲۷۲۰۰۰, ۲۷۲۵۰۰, ۲۷۳۰۰۰, ۲۷۳۵۰۰, ۲۷۴۰۰۰, ۲۷۴۵۰۰, ۲۷۵۰۰۰, ۲۷۵۵۰۰, ۲۷۶۰۰۰, ۲۷۶۵۰۰, ۲۷۷۰۰۰, ۲۷۷۵۰۰, ۲۷۸۰۰۰, ۲۷۸۵۰۰, ۲۷۹۰۰۰, ۲۷۹۵۰۰, ۲۸۰۰۰۰, ۲۸۰۵۰۰, ۲۸۱۰۰۰, ۲۸۱۵۰۰, ۲۸۲۰۰۰, ۲۸۲۵۰۰, ۲۸۳۰۰۰, ۲۸۳۵۰۰, ۲۸۴۰۰۰,

- of Teams of Learning Automata", in Symp. Intell. Syst., Bangalore, Dec. 1991.
- [8] M. A. L. Thatachar and V. V. Phananskar, "Learning Global Maximum with Parameterized Learning Automata", IEEE Trans. Neural Networks, Vol. 6, No. 2, March 1995.
- [9] Kumpan S. Narendra and M. A. L. Thatachar, Learning Automata an Introduction, New Jersey, Prentice Hall, 1989.
- [10] K. Najim and A. S. Poznyak, Learning Automata Theory and Applications, Elsevier Science Ltd, 1994.
- [11] P. Mars, J.R. Chen and R. Nombiar, Learning Algorithms Theory and Applications in Signal Processing, Control and Communications, CRC Press, New York, 1998.
- [12] S. Lakshimavarahan, Learning Algorithms Theory and Applications, New York, Springer - Verlag, 1981.
- [13] J. T. Tou and R. C. Gonzalez, Pattern Recognition Principles, Reading, MA : Addison-Wesley, 1974.
- [14] M. R. Meybodi and S. Lakshimavarahan, "On a class of Learning Algorithms which have a symmetric Behavior under Success and Failure", Springer - Verlag Lecture Notes in Statistics, PP. 145-155, 1984.
- [15] M.R. Meybodi, 'Results on Strongly Absolutely Expedient Learning Automata", Proceedings of Inference Conference 86, ed. D.R. Moates and R. Butrick (Athens, Ohio : Ohio University Press, 1987), pp. 197-209.

جدول ۱

ردیف	روش	پارامترها	وقت	تعداد مرحله
۱۰۸۷۴	Sastry , Thatachar	$\lambda_1 = 0.001$	۹۵/۹۵۰	
۲۶۶۲۱	$L_{R+1}$	$a = 0.001$	۹۷/۹۱۶	
۵۰۰۱۲	$L_{R+2P}$	$a = 0.002$ $b = 0.0001$	۹۷/۹۵۰	
۱۰۰۹۶	TSetline $L_{KN,K}$	$N = 1$ $W = 1..$	۹۸/۹۸۶	
۱۲۰۱۱	Krinsky	$N = 1$ $W = e.$	۹۸/۹۷۴	برای عمل حافظه ۱
۷۸۹۷۶	TSetline $L_{KN,K}$	$N = 2$ $W = 10..$	۹۸/۹۰۱	
۷۷۱۴۶	Tsetline $G_{KN,K}$	$N = 2$ $W = 10..$	۹۸/۹۷۹	
۷۵۹۷۷	Krinsky	$N = 2$ $W = 15..$	۹۸/۹۸۱	
۷۵۹۷۸	Krylov	$N = 2$ $W = 15..$	۹۸/۹۹	
۷۸۹۹۸	آنلماتون نوع ۱	$W = 13..$	۹۸/۹۹۰	
۷۸۹۷۵	آنلماتون نوع ۲	$W = 18..$	۹۸/۹۸۱	
۷۸۹۷۷	آنلماتون نوع ۳	$W = 11..$	۹۸/۹۷۷	
۷۸۹۷۸	آنلماتون نوع ۴	$W = 50..$	۹۸/۹۷۸	
۷۸۹۷۹	آنلماتون نوع ۵	$N_a = 70..$	۹۸/۹۷۹	
۷۸۹۷۰	آنلماتون نوع ۶	$N_a = 12..$	۹۸/۹۸۰	
۷۸۹۷۱	قانون بیز	$P(W_1) = \frac{1}{2}$ $P(W_2) = \frac{1}{2}$	۹۸/۹۸۱	

نتایج بدست آمد: نشان دادند که آنلماتونهای پادگیر با ساختار ثابت با عمل حافظه ۱ ، از سرعت همگرایی و دقت بهتری برخوردار می‌باشد. در بخش ۸ با مصلاح آنلماتون ساختار ثابت با عمل حافظه ۱ ، آنلماتونهای ساختار ثابت جدیدی با کارایی بالاتر ارائه شدند. در بخش ۹ آنلماتونهای ساختار ثابت دیگری پیشنهاد گردید که دارای عمل حافظه بزرگی بیشتر و طریق تراپر فاکل شدن بین پاسخ محیط و تبیه و شویت آنلماتونها کاربری بالاتر از تدازه گیری کرد. طبق قانون بیز برای جدا کردن کلاس  $W_1$  از  $W_2$ ، معادله زیر می‌باشد:

$$h_1(x) = p(W_1)f(x|W_1) - p(W_2)f(x|W_2) \quad (23)$$

که در آن  $f(x|W_i)$  تابع جگالی شرطی  $p(W_i)$  احتمال انتخاب کلاس  $W_i$  است. در مسنداده از رابطه (23) برای جذلزاری کلاس  $W_1$  از دو کلاس دیگر، کلاسهاي  $W_1, W_2, W_3$  را به صورت زیر در نظر می‌گیریم :

$$W_2 = W_1 + W_3, \quad W_1 = W_2$$

احتمال انتخاب کلاس  $W_1$  و  $p(W_1)$  احتمال انتخاب هر یک از دو کلاس  $W_2$  یا  $W_3$  می‌باشد. از آنجا که در بیاده‌سازی الگوریتم‌های پادگیری در بحث‌های گذشته در مرحله تابع تبايز روی یک نمونه تصادفی از مجموعه کل نمونه‌ها آزمایش شده است، پس احتمال تعلق دلشن نمونه به کلاس  $W_1$  برای  $\frac{1}{2}$  احتمال تعلق داشتن نمونه به کلاسهاي  $W_2$  یا  $W_3$  برابر  $\frac{1}{2}$  خواهد بود و در این صورت  $h_1(x) = h_2(x)$ .  $p(W_1) = \frac{1}{2}, p(W_2) = \frac{1}{2}$  برای جذلزاری کلاسهاي  $W_1, W_2$  از سایر کلاسها به طور مشابه پیاده‌سازی می‌شوند. نتیجه ۱۰۰٪ بار آزمایش قانون بیز روی ۱۰۰٪ مجموعه نمونه مختلف، دقت متوسط  $98/974$  بدست می‌دهد (برای  $\frac{1}{2}$ )  $= p(W_1) = p(W_2) = \frac{1}{2}$  ، دقت  $95/107$  حاصل می‌شود. جدول شماره ۱ حلاجنه نتایج بدست آمده از کارایی روش‌های مختلف ارائه شده در این مقاله در حل مسئله نمونه را نشان می‌دهد.

## ۱۱ - مراجع

- [1] Fukunaga, An Introduction to Statistical Pattern Recognition, New York, Academic Press, 1972.
- [2] Pierre A. Devijver and Joseph Kittler, Pattern Recognition Theory and Applications, NATO ASI Series, Series F : Vol. 30, 1986.
- [3] Andrew G. Barto and P. Anandan, "Pattern Recognition Stochastic Learning Automats", IEEE Trans. Sys. Man and Cybern., Vol. SMC-15, No. 3, May/January 1985.
- [4] Mandayam A. L. Thatachar and P. S. Sastry, "Learning Optimal Discriminant Functions Through a Cooperative Game of Automata", IEEE Trans. Sys. Man and Cybern., Vol. SMC-17, No. 1, January/February 1987.
- [5] M. A. L. Thatachar and V. V. Phananskar, "A feedforward Network of Learning Automata for Pattern Recognition", in Proc. Int. Joint Conf. Neural Networks, Singapore, Nov. 1991.
- [6] M. A. L. Thatachar and V. V. Phananskar, "Convergence of Teams and Hierarchies of Learning Automata in Connectionist Systems", IEEE Trans. Sys. Man and Cybern., Vol. 25, No 11, Nov. 1995.
- [7] V. V. Phananskar and M. A. L. Thatachar, "Global Convergence