



دانشکده صنعتی امیرکبیر
(پلی تکنیک تهران)
دانشکده مهندسی کامپیوتر

گزارش درس روش پژوهش و ارائه

بررسی الگوریتم تکامل عصبی عمیق در آموزش شبکه های عصبی

نگارش

کیوان ایچی حق

استاد مشاور

دکتر رضا صفا بخش

بهار ۱۴۰۱

سپاس‌گزاری

در ابتدا لازم میدانم که از استاد گرانقدر دکتر رضا صفا بخش نهایت تشکر را داشته باشم که مطالب بسیار مفیدی را به من آموختند و تهیه این گزارش بدون راهنمایی‌های ایشان امکان‌پذیر نبود.

کیوان ایپچی حق

بهار ۱۴۰۱

چکیده

احساس نیاز به سیستم های خبره برای کمک در انجام مسائل مختلف به سرعت در حال افزایش است. امروزه بیشتر از قبل در مسائل روزمره به هوش مصنوعی وابسته بوده و با افزایش تقاضا و همچنین افزایش تنوع مسئله های مطرح شده، نیاز به سیستم هایی داشته که با خطای کمتر و در زمان کمتر خروجی مطلوب تری بدهند.

در گذشته شبکه های عصبی طراحی شده بیشتر بر روی بازی ها و بر اساس میزان دستیابی به هدف از پیش تعریف شده آزمایش می شدند، اما با گذشت زمان شبکه های پیشرفته تری طراحی شده اند که قادر هستند خود به تعریف درستی از هدف برسند و دخالت انسان را در روند الگوریتم کمتر کنند. از طرفی روش های مدرنی بر روی کار آمدند که قادر بودند در بسیاری از شرایط مسئله از روش های سنتی پیشی بگیرند.

محدوده این گزارش بر حوزه استراتژی های تکاملی عمیق است که با معرفی الگوریتم های برتر و تحلیلی ویژگی های آنها سعی بر بهبود کارایی و افزایش قدرت محاسباتی و دقت این الگوریتم ها هستیم.

در ابتدا سعی بر شناخت الگوریتم های پایه یادگیری عمیق داشته و بعد از معرفی چندین ویژگی بارز این الگوریتم با چندین روش مطرح آن آشنا می شویم. سپس الگوریتم تکامل عصبی عمیق را معرفی و جنبه های مختلف آنرا بررسی می کنیم. در فصل بعد با استفاده از آموخته های دو فصل پیشین ساختار ها و ویژگی های مختلف این الگوریتم را بررسی کرده و به نتایج قابل ملاحظه ای رسیده و نتیجه می گیریم هر الگوریتم لزوما برای همه مسائل بهینه عمل نمیکند. چراکه بر مشکل نیازمند راه حل و در نتیجه الگوریتم مختص به خودش است و باید طراحی الگوریتم نسبت به مسئله طراحی شود؛ اما به طور کلی می توان از الگوریتم های خلاقانه برای حل مسائلی که بیشتر نیاز به خلاقیت دارند استفاده کرده و از الگوریتم های هدفمند برای مسائلی که می خواهیم بدون توجه به مسیر، به هدف مشخص شده برسیم.

در انتها نتیجه گیری مبنی در عملکرد الگوریتم های مختلف داشته و بررسی می کنیم این الگوریتم ها در چه شرایطی بازدهی بالاتر داشته و مورد نیاز مسئله هستند.

واژه های کلیدی:

استراتژی های تکاملی، الگوریتم ژنتیک، تابع هدف، الگوریتم جستجو تازگی، الگوریتم تکامل عصبی عمیق، الگوریتم نیت¹، شبیه سازی دوپا، مسئله مارپیچ

¹ Neat (Neuroevolution of Augmenting Topologies)

فهرست مطالب

عنوان	صفحه
فصل اول.....	۱
۱- مقدمه.....	۲
فصل دوم.....	۳
۲- استراتژی های تکاملی.....	۳
۱-۲ ویژگی ها	۴
۱-۱-۲ مقیاس پذیری و موازی سازی.....	۵
۲-۲-۲ پارامتر سازی شبکه.....	۷
۲-۲ روش ها	۷
۲-۲-۲ الگوریتم ژنتیک.....	۷
۲-۲-۲ جست و جو تازگی.....	۹
فصل سوم.....	۱۰
۳-۴ تکامل عصبی عمیق.....	۱۰
۱-۳ رمز گذاری	۱۱
۱-۱-۳ روش غیر مستقیم.....	۱۱
۲-۱-۳ روش غیر مستقیم.....	۱۱
۳-۴ جهش	۱۲
۴-۳ معاونت های رقابتی.....	۱۳
۳-۴ گونه سازی.....	۱۴
فصل چهارم.....	۱۵
۴- استفاده از الگوریتم نیت در حل مسائل	۱۵
۱-۴ شبکه مارپیچ فریبده.....	۱۶
۲-۴ آزمایش دوبا انسان نما.....	۱۷
3-4 آزمایش بازی آتاری.....	۱۹
فصل پنجم.....	۲۰
۵- نتیجه گیری و پیشنهاد ها.....	۲۰
۱-۵ نتیجه گیری	۲۱
۲-۵ پیشنهاد ها	۲۱

فهرست اشکال

عنوان	صفحه
شکل ۱- شبه کد استراتژی های تکاملی.....	۴
شکل ۲- شبه کد استراتژی های تکاملی موازی سازی شده.....	۵
شکل ۳- تاثیر بروز رسانی پارامتر های کارگران بر امتیاز بدست آورده شده.....	۶
شکل ۴- تاثیر افزایش تعداد کارگران موازی سازی شده بر روی زمان انجام محاسبات.....	۶
شکل ۵- فرمول رمز گذاری جدید برای اعضای جمعیت.....	۷
شکل ۶- فرایند تکرار شونده تولید فرزند توسط ۲ والد در الگوریتم ژنتیک.....	۸
شکل ۷- روش های مختلف برای جهش در نقاط مختلف رشته ژن ها.....	۱۲
شکل ۸- از دست رفتن اطلاعات در دو شبکه که به گره مرکزی وابسته هستند.....	۱۲
شکل ۹- استفاده از نشانه های تاریخی برای همسو سازی آسان تر والد ها.....	۱۳
شکل ۱۰- مقایسه الگوریتم ها در حل مسئله مارپیچ.....	۱۴
شکل ۱۱- مقایسه بین دو الگوریتم در شبیه سازی واقعی.....	۱۶
شکل ۱۲- مقایسه دو الگوریتم در شبیه سازی مارپیچ.....	۱۷
شکل ۱۳- شکل های مارپیچ بعد از ۲۵۰ هزار نسل.....	۱۷
شکل ۱۴- یک دوبا که توسط یک شبکه عصبی قادر به ایستادن است.....	۱۸
شکل ۱۵- مقایسه عملکرد هر دو الگوریتم در ۵۰ تست متوالی.....	۱۸
شکل ۱۶- مقایسه امتیاز ثبت شده توسط الگوریتم های مختلف یادگیری عمیق.....	۱۹

فصل اول

مقدمه

۱- مقدمه

در زمینه یادگیری عمیق، شبکه های عصبی عمیق با لایه های بسیار زیاد و میلیون ها نورون و اتصالات اکنون به طور معمول از طریق شیب نزولی گرادیان^۲ آموزش داده می شوند. این روش در اکثر مسائل به خوبی عمل میکند اما هیچ روشی بدون ایراد نیست؛ وجود پارامترهای زیاد در روند آموزش شبکه ها و دخالت انسان در تنظیم این پارامترها از نکات منفی این روش هستند. از دیگر مشکلات این الگوریتم بروز شدن اجباری تمام نورون ها در هر چرخه است؛ در صورتی که برخی نورون ها، لزوماً نیاز به بروز شدن ندارند. با این حال، در این مقاله بررسی خواهیم کرد که تکامل عصبی عمیق^۳ که در آن شبکه های عصبی از طریق الگوریتم های استراتژی تکاملی بهینه سازی می شوند هم روشی مؤثر برای آموزش شبکه های عصبی عمیق در حل مسائل یادگیری تقویتی^۴ کاربرد دارد.

هدف این گزارش معرفی روش های تکاملی به عنوان جایگزینی برای روش های سنتی آموزش شبکه های عصبی بوده و سپس مقایسه این روش ها با یکدیگر برای انتخاب درست الگوریتم بسته به شرایط مسئله ای که قصد حل کردن آن را داریم.

در فصل دوم ابتدا به بررسی استراتژی های تکاملی پرداخته و ویژگی های کلیدی آن که موجب ایجاد تمایز در مقابل الگوریتم ها و استراتژی های دیگر شده را مورد مطالعه قرار می دهیم. سپس دو الگوریتم ژنتیک و جست و جو تازگی را مورد بررسی قرار داده و با یکدیگر مقایسه می کنیم.

در فصل سوم با کمک آموزه های فصل پیشین الگوریتم های تکاملی عصبی عمیق را معرفی کرده و معروف ترین آنها به نام نیت و ویژگی های مهم آن را که موجب انقلابی بودن این الگوریتم شده است را مورد بررسی قرار می دهیم.

در فصل آخر با استفاده از هر سه الگوریتم مطرح شده چندین آزمایش و شبیه سازی مانند انسان نما دو پا و مارپیچ، تحلیل کرده و الگوریتم ها را با یکدیگر مقایسه می کنیم.

از محدودیت های این پژوهش، نیازمندی به سخت افزار قدرتمند برای اجرای الگوریتم های تکاملی بوده که باعث شده نتوان برخی شبیه سازی ها را در دستگاه های معمولی انجام داد. همچنین فرض می کنیم مسائلی که به دنبال حل آنها هستیم نسبتاً ساده بوده و وارد جزئیات پیاده سازی و حالت های خاص آنها نمی شویم. برای آماده سازی این گزارش الگوریتم های مطرح شده بر روی مسئله ساده هدایت موشک از مبدا به مقصد و همچنین بازی پرنده فلاپی به جهت یادگیری طراحی و پیاده سازی شد.

^۲ Stochastic Gradient Decent (SGD)

^۳ Neuroevolution

^۴ Reinforcement Learning (RL)

فصل دوم

استراتژی های تکاملی

۲- استراتژی های تکاملی^۵

استراتژی های تکاملی یک کلاس از الگوریتم های بهینه سازی در سری الگوریتم های جست و جوی اکتشافی است که از تکامل طبیعی الهام گرفته شده و به آن ها همانند شبکه های عصبی، به صورت یک جعبه سیاه نگاه می شود. این استراتژی ها متعلق به الگوریتم های تکاملی بوده که از جهش، ترکیب مجدد و انتخاب کاندید ها برای چرخه بعدی که اصطلاحاً به آن نسل هم گفته می شود، استفاده می کنند. برای درک بهتر این موضوع شبه کد [1] کلی این استراتژی ها آورده شده است:

Algorithm 1 Evolution Strategies

```

1: Input: Learning rate  $\alpha$ , noise standard deviation  $\sigma$ , initial policy parameters  $\theta_0$ 
2: for  $t = 0, 1, 2, \dots$  do
3:   Sample  $\epsilon_1, \dots, \epsilon_n \sim \mathcal{N}(0, I)$ 
4:   Compute returns  $F_i = F(\theta_t + \sigma \epsilon_i)$  for  $i = 1, \dots, n$ 
5:   Set  $\theta_{t+1} \leftarrow \theta_t + \alpha \frac{1}{n\sigma} \sum_{i=1}^n F_i \epsilon_i$ 
6: end for

```

شکل ۱- شبه کد استراتژی های تکاملی

الگوریتم های داخل این کلاس از نظر نحوه نمایش جمعیت، اعمال جهش و ترکیب با هم تفاوت هایی دارند. از معروف ترین اعضای این کلاس می توان به استراتژی تکامل انطباق ماتریس کوواریانس^۶ و الگوریتم ژنتیک^۷ اشاره کرد.

الگوریتم های بهینه سازی جعبه سیاه مانند استراتژی های تکاملی علاوه بر موازی سازی آسان و داشتن مزیت در موارد متعددی با توالی های اکشن طولانی و پاداش های تاخیری، مزایای دیگری نسبت به تکنیک های سنتی دارند که گرادیان ها را محاسبه می کنند. برای مثال این استراتژی ها می توانند با پاداش های کم و تأخیر محاسبه و اعمال پاداش ها مقابله کند، زیرا تنها بازده کلی یک قسمت استفاده می شود، در حالی که روش های دیگر از پاداش های فردی به جای گروهی استفاده می کنند [2].

از طرف دیگر با عدم نیاز به محاسبه انتشار پس زمینه^۸، بهینه سازهای جعبه سیاه مقدار محاسبات در هر قسمت را حدود ۶۶٪ و حافظه را به طور بالقوه به مقدار زیادی کاهش می دهند. علاوه بر این، محاسبه نکردن صریح یک گرادیان از مشکلات ناشی از گرادیان های انفجاری محافظت می کند که هنگام کار با شبکه های عصبی مکرر رایج هستند، به طوری که دیگر نیاز به تکنیک های متعددی که استفاده می شود تا مانع این رخداد ها شود نداریم. در نهایت، این استراتژی ها به ما اجازه می دهد تا عناصر غیرقابل تمایز را در معماری خود بگنجانیم و به سادگی و با انعطاف پذیری بالا آنها را تغییر دهیم.

در قدم بعد، ابتدا به ویژگی های مهم این استراتژی ها که موجب تمایز آنها با الگوریتم های دیگر شده می پردازیم:

⁵ Evolution Strategies (ES)⁶ Covariance Matrix adaptation Evolution Strategy (CMA-ES)⁷ Genetic Algorithm (GA)⁸ Backpropagation

۱-۲ ویژگی ها

۱-۱-۲ مقیاس پذیری و موازی سازی

از مزایای قابل توجه این کلاس از استراتژی ها، مقیاس پذیر بودن آنها در استفاده از تعداد زیادی کارگران به صورت موازی است. زیرا هر کارگر قادر است بدون وابستگی به دیگر کارگران کار خود را انجام داده و در نهایت خروجی مطلوب را باز گرداند؛ زیرا اگر آشفتگی دانه های تصادفی^۹ را از پیش بین کارگران همگان سازی کرده باشیم تا تغییر آشفتگی به صورت موازی و باهم باشد، هر کارگر تنها با فرستادن و دریافت یک عدد می تواند پارامتر های خود را بروز رسانی کند که باعث از بین رفتن بیشتر وابستگی بین کارگران می شود. این ویژگی در تمایز با روش های گرادیان است که در آن کارگران باید به همواره با یکدیگر در ارتباط باشند زیرا خروجی کارگران به یکدیگر وابسته اند.

در شکل زیر، شبه کدی را برای پیاده سازی استراتژی های تکاملی موازی سازی شده را میبینیم که به استفاده از دانه های تصادفی مشترک استفاده میکنند [1]:

Algorithm 2 Parallelized Evolution Strategies

```

1: Input: Learning rate  $\alpha$ , noise standard deviation  $\sigma$ , initial policy parameters  $\theta_0$ 
2: Initialize:  $n$  workers with known random seeds, and initial parameters  $\theta_0$ 
3: for  $t = 0, 1, 2, \dots$  do
4:   for each worker  $i = 1, \dots, n$  do
5:     Sample  $\epsilon_i \sim \mathcal{N}(0, I)$ 
6:     Compute returns  $F_i = F(\theta_t + \sigma \epsilon_i)$ 
7:   end for
8:   Send all scalar returns  $F_i$  from each worker to every other worker
9:   for each worker  $i = 1, \dots, n$  do
10:    Reconstruct all perturbations  $\epsilon_j$  for  $j = 1, \dots, n$  using known random seeds
11:    Set  $\theta_{t+1} \leftarrow \theta_t + \alpha \frac{1}{n\sigma} \sum_{j=1}^n F_j \epsilon_j$ 
12:   end for
13: end for

```

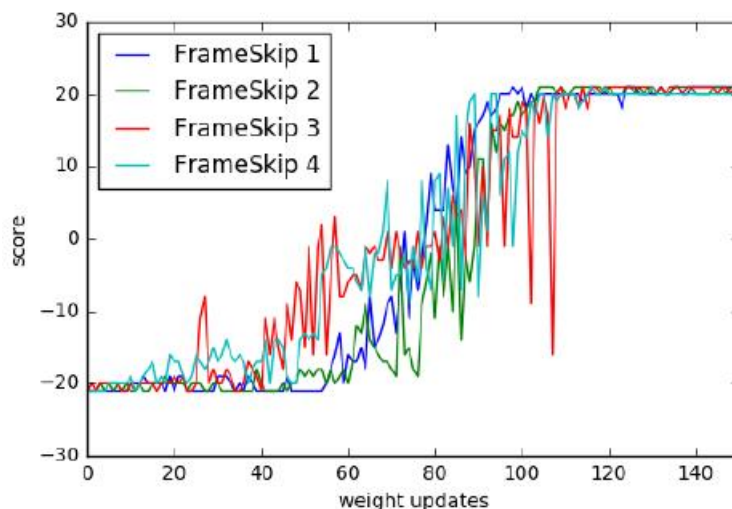
شکل ۲- شبه کد استراتژی های تکاملی موازی سازی شده

در عمل، نمونه گیری را به این صورت اجرا می کنیم که هر کارگر ابتدا جمعیت مرتب شده ای را با استفاده از نویز گاوسی^{۱۰} (توزیع نرمال) واحد بین تمام کارگران ساخته و سپس پارامتر ها را با اضافه کردن تصادفی نویز در هر تکرار، آشفته می کنیم. اگرچه این بدان معنی است که اغتشاشات در سراسر تکرارها کاملاً مستقل از یکدیگر نیستند، مشکل خاصی در عملکرد الگوریتم بوجود نمی آید. (آشفتگی ها لزوماً مستقل از یکدیگر نیستند زیرا می توان دو کارگری داشت که آشفتگی یکسانی از فضای حالت داشته باشند اما این احتمال در فضای نمونه به اندازه بزرگ به صفر میل می کند)

⁹ Random Seeds

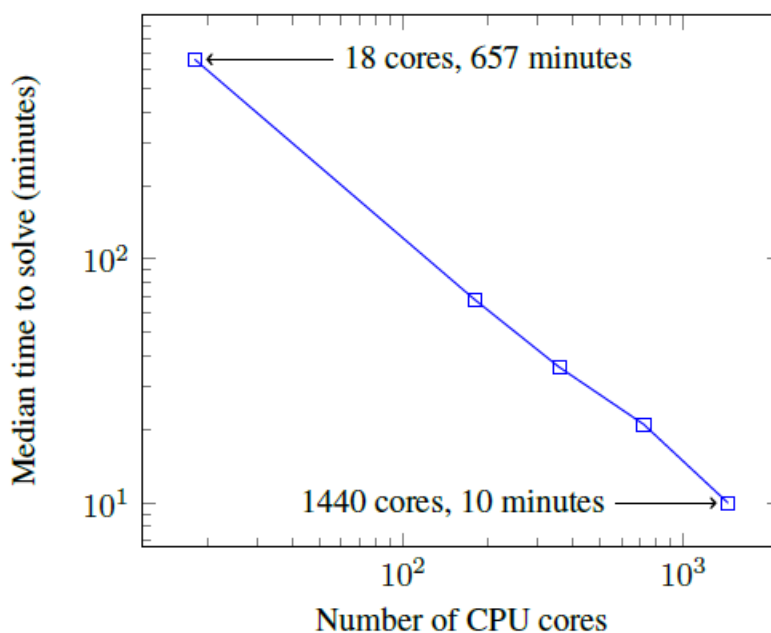
¹⁰ Gaussian Distribution (Normal Distribution)

فصل دوم – استراتژی های تکاملی



شکل ۳- تاثیر بروز رسانی پارامتر های کارگران بر امتیاز بدست آورده شده

همانطور که بالاتر مشاهده می شود کارگران موازی سازی شده در انتهای فرایند خود پارامتر ها و نتایج بدست آمده شده را با یکدیگر به اشتراک گذاشته که باعث شده عملکرد کلی الگوریتم افزایش قابل ملاحظه ای داشته باشد. همچنین عملکرد هر کارگرد به تنهایی مطلوب نیست، اما در صورت کارکرد گروهی عملکرد کلی افزایش خواهد داشت [1]:



شکل ۴- تاثیر افزایش تعداد کارگران موازی سازی شده بر روی زمان انجام محاسبات

نکته حائز اهمیت دیگر آن است که با افزایش تعداد کارگران تا وقتی که مدت زمان مورد نیاز برای ساختن و مقدار دهی کردن اولیه کارگران کمتر از مدت زمان اجرا شدن هر کارگر به تنهایی باشد، زمان کلی اجرای الگوریتم به شکل قابل توجهی کاهش می

یابد. همانطور که در شکل پیدا است الگوریتم با استفاده از ۱۸ هسته موازی در ۶۵۷ دقیقه (حدوداً ۱۱ ساعت) اجرا شده اما با افزایش تعداد کارگران موازی شده به ۱۴۴۰ عدد، مدت زمان به ۱۰ دقیقه کاهش یافته است.

۲-۲-۲ پارامتر سازی شبکه

همانطور که در بخش قبلی اشاره شد، استراتژی های تکاملی با استفاده از اغتشاش پارامتر ها عمل می کنند. زمانی این استراتژی عملکرد قابل قبولی خواهد داشت که بخشی از جمعیت نسبت به بقیه عملکرد بهتری از خود نشان دهند؛ پس حیاتی است که توزیع نرمال به صورت تصادفی، پارامتر ها و در نتیجه جمعیت بهتری برای نسل های آینده تولید کند [3].

نرمال سازی دسته مجازی^{۱۱} یک تکنیک استفاده شده برای آموزش در این استراتژی است که همانند نرمال سازی عادی دسته ای عمل میکند. در این تکنیک یک دسته برای شروع آموزش انتخاب شده و در طی یک آموزش ثابت می ماند. این عمل سبب می شود که مدل در مراحل اولیه آموزش نسبت به تغییرات کوچک و ناگهانی واکنش بیشتری بدهد و به همین علت نیاز است حجم جمعیت در هر دسته را به اندازه کافی بزرگ باشد تا مانع تاثیر زیاد تغییرات کوچک اطلاعات بر عملکرد کلی مدل شویم. این تکنیک برای مدل های پیچیده و گسترده هزینه زیادی به همراه دارد، اما در مدل های نسبتاً ساده این هزینه قابل چشم پوشی است و باعث بهبود فرایند یادگیری مدل می شود [3].

الگوریتم ژنتیک که کمی جلو تر به بررسی دقیق تر آن می پردازیم، به طور سنتی هر عضو جمعیت را به صورت بردار θ ذخیره می کند که باعث می شود از نظر حافظه بهینه نباشد. از طرفی مقیاس پذیری پایینی دارد و اندازه شبکه عصبی را به صورت چشمگیری افزایش میدهد. برای حل این مشکل از روش جدیدی استفاده می کنیم که در آن هر بردار پارامتر را با استفاده از دانه اولیه^{۱۲} و یک لیست از دانه های تصادفی که دنباله ای از جهش هایی میسازند که منجر به ساخته شدن هر θ می شود [4]:

$$\theta^n = F(\theta^{n-1}, T_n) = \theta^{n-1} + \sigma \varepsilon(T_n)$$

شکل ۵- فرمول رمز گذاری جدید برای اعضای جمعیت

در این فرمول T_n مجموعه ای از دانه های هر جهش بوده و $\varepsilon(T_n) = N(0,1)$ یک تولید کننده اعداد تصادفی با توزیع گاوسی است. با استفاده از این تکنیک میتوان شبکه عصبی با بیش از ۴ میلیون پارامتر را در تنها چندین هزار بایت ذخیره کرد. برای افزایش اهمیت این موضوع می توان به این نکته اشاره کرد که آخرین مدل $GPT-3$ دارای ۱۷۵ میلیارد پارامتر است که حدوداً ۴۵ ترا بایت حجم میگیرد.

۲-۲ روش ها

۲-۲-۲ الگوریتم ژنتیک^{۱۳}

هر چرخه از این الگوریتم ابتدا با تشکیل جمعیت شروع شده و پس از یک جهش تصادفی که برای طبیعی نگه داشتن روند تکامل حیاتی است، مقدار تابع هدف که عملاً امتیاز اعضای جدید است، ارزیابی میشود. در نهایت بخشی از جمعیت که

¹¹ Virtual Batch Normalization

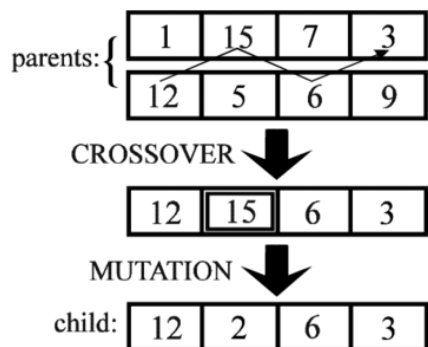
¹² Initialization Seed

¹³ Genetic Algorithm

فصل دوم – استراتژی های تکاملی

بالاترین امتیاز را دارند ترکیب شده تا جمعیت جدیدی ساخته شود. این روند تا جایی ادامه دارد که به طور کامل جمعیت تکامل یافته و به هدف نهایی خود برسند. حال با ثابت نگه داشتن روند کلی و تغییر دادن کارکرد اجزا آن میتوان نمایش های مختلفی بدست آورد.

شکل ۴ مثال ساده ای از روند این چرخه را برای دو والد انتخابی از چرخه پیشین نشان میدهد. هر عضو این جمعیت ژن خود را دارد که توسط تابع ترکیب شده و در نهایت تعدادی ژن در مرحله جهش به تصادف جهش پیدا می کنند. جهش عضو مهمی از این چرخه است زیرا از به دام افتادن الگوریتم در نقاط بحرانی محلی جلوگیری میکند.



شکل ۴- فرایند تکرار شونده تولید فرزند توسط ۲ والد در الگوریتم ژنتیک [1]

در ابتدا با استفاده از یک الگوریتم ژنتیک بسیار ساده آزمایشی را پیش برده تا نحوه عملکرد یک الگوریتم ژنتیک بدون گرادیان را مورد مطالعه قرار دهیم. در این الگوریتم برای تابع ترکیب از روش برش و انتخاب^{۱۴} استفاده کردیم که در آن درصدی از جمعیت با بالاترین امتیاز بعد از اعمال جهش تصادفی با الگوی نوین گاوسی برای تولید فرزند انتخاب میشوند.

برای درک عمیق تر دلیل انقلابی بودن این الگوریتم در زمان خود، یک مسئله بهینه سازی مطرح میکنیم: تصور کنید در این مسئله ۶ پارامتر داریم که هر یک از آنها می توانند ۶ حالت مختلف داشته باشند. اگر فرض کنیم ارزیابی هر ترکیب ۶ تایی پارامتر ها و محاسبه امتیاز این ترکیب یک دقیقه زمان نیاز داشته باشد، با روش بی رحمانه به ۶ دقیقه (معادل ۳۲ روز) زمان نیاز داریم؛ اما اگر روشی داشتیم که در هر نسل فقط ۶۰ عضو برتر را آزمایش می کرد و ۶ نسل داشتیم، به ۶۰۰ دقیقه (معادل ۱۰ ساعت) نیاز خواهیم داشت [5].

درست است که این یک تخمین تقریبی است و همیشه ۶۰ عضو برتر انتخاب نمی شوند، ولی با این وجود می توانید از هزینه های سربار ای که می توان از آنها جلوگیری کرد شهود خوبی پیدا کنید. دقت این الگوریتم به تابع هدف، تعداد نسل و پارامتر های دیگر بستگی دارد اما با استناد به نتایج به دست آمده می توان آن را یک الگوریتم اکتشافی موفق بنامیم.

بنابراین الگوریتم به گونه ای طراحی شده است که برای یک تابع هدف از پیش تعریف شده و تعداد نسل هایی که باید تکرار شوند اجرا شود و جمعیت هر نسل نیز مشخص شده است. بنابراین اعضای با امتیاز کمتر در هر نسل از بین می روند، اما آیا این تنها روشی است که می توانیم با استفاده از آن در زمان کمتری به بهترین راه حل های ممکن برسیم؟ آیا لزومی دارد اعضای ناکارآمد

¹⁴ Truncation Selection

جمعیت از بین بروند؟ آیا دخالت انسان در تنظیم توابع هدف می تواند بر عملکرد الگوریتم ها تاثیر منفی بگذارد؟ در ادامه الگوریتم جایگزین دیگری را بررسی میکنیم که بخشی از مشکلات مطرح شده را حل می کند:

۲-۲-۲ جست و جو تازگی^{۱۵}

هدف نهایی از هوشمند سازی الگوریتم ها آن است که پارامتر هایی مانند آنچه در الگوریتم ژنتیک به صورت دستی تنظیم می کردیم نیز توسط خود الگوریتم و با توجه به شرایط تنظیم شوند یا به عبارتی پویا باشند. الگوریتم جست و جو تازگی سعی بر حل این مشکل دارد به گونه ای که این الگوریتم به پیشرفتی که توسط اهداف یا عملکرد تعریف شده است پاداش نمی دهد، بلکه به متفاوت بودن پاداش می دهد. به عبارتی تکامل آن بدون هدف والا است [2].

هدف از طراحی این الگوریتم آن است که رشد پیچیدگی در تکامل طبیعی، محصول جانبی بهینه سازی تابع هدف نیست، بلکه نتیجه تمایل بی پایان آن برای کشف راه های مختلف برای رویارویی با چالش های زندگی است. این ویژگی منجر به کشف دائمی تازگی و افزایش پیچیدگی مدل می شود.

برای بهتر نشان دادن این موضوع و مقایسه دو الگوریتم مطرح شده – الگوریتم ژنتیک و الگوریتم جست و جو تازگی – می توان آزمایش شبیه سازی حرکت دوپا را مثال زد که در آن تابع هدف الگوریتم ژنتیک ممکن است خم شدن دوپا را در نسل اول به عنوان یک حرکت نامطلوب از بین ببرد، اما الگوریتم جست و جو تازگی به این تفاوت و خلاقیت در هر نسل بعدی تا زمانی که پایداری با مقدار خمش مناسب حاصل شود، پاداش خواهد داد.

¹⁵ Novelty Search

فصل سوم

تكامل عصبي عميق

۴-۳ تکامل عصبی عمیق^{۱۶}

استراتژی تکامل عصبی عمیق راه کار جایگزین برای آموزش، بهبود و تکامل شبکه های عصبی است که اولین و معروف ترین الگوریتم آن نیت، این موضوع را به عنوان یک رویکرد قابل اجرا نشان داد.

این الگوریتم که فقط از شبکه نود^{۱۷} های متراکم تشکیل می شود، قابلیت آن را می دهد که بتوانیم از الگوریتم های مختلف مانند الگوریتم ژنتیک و الگوریتم جست و جو تازگی و همچنین از مفاهیم مختلف مانند تکامل و یا کاهش شیب نزولی گرادیان استفاده کنیم و یا حتی از ترکیب همه این مفاهیم. مقالات اخیر حتی راه هایی را برای استفاده از الگوریتم های نیت و مانند آن برای ساختار شبکه عصبی تکامل یافته و سپس استفاده از کاهش شیب نزولی گرادیان و حرکت عقب گرد برای بهینه سازی این شبکه ها برجسته کرده اند، منطقه ای که در آینده نزدیک بی اندازه مرتبط و مهم خواهد شد.

این الگوریتم راه حل مناسبی برای مشکلات زیادی که پیش از آن وجود داشت ارائه داد که به مهم ترین آنها می پردازیم:

۱-۳ رمز گذاری

در زیست شناسی ما یک ژنوتیپ^{۱۸} و یک فنوتیپ^{۱۹} داریم. ژنوتیپ بازنمایی ژنتیکی یک موجود و فنوتیپ نمایش فیزیکی واقعی موجود است. استراتژی ها و الگوریتم های تکاملی همیشه آینده زیست شناسی هستند و تکامل عصبی هم از این نظر تفاوتی ندارد. برای رمز گذاری ژن های طراحی شده مسئله دو روش مستقیم و غیر مستقیم داریم:

۱-۱-۳ روش غیر مستقیم

یک رمز گذاری مستقیم به صراحت همه چیز را در مورد یک فرد مشخص می کند. اگر یک شبکه عصبی داشته باشیم، به این معنی است که هر ژن مستقیماً به گره، اتصال یا ویژگی شبکه مرتبط می شود. این می تواند رمز گذاری ۱ و ۰ باشد، رمز گذاری گراف (به هم پیوستن گره های مختلف با اتصالات وزنی) و یا حتی روشی پیچیده تر باشد. نکته آن است که همیشه ارتباط مستقیمی بین ژنوتیپ و فنوتیپ وجود خواهد داشت که بسیار واضح و خواندنی است.

۲-۱-۳ روش غیر مستقیم

رمز گذاری غیر مستقیم دقیقاً برعکس روش مستقیم است. به جای آنکه مستقیماً مشخص شود یک ساختار چگونه تعریف شود، رمز گذاری های غیرمستقیم تمایل دارند قوانین یا پارامترهای فرآیندها را برای ایجاد یک ژن مشخص کنند. در نتیجه، رمز گذاری های غیر مستقیم بسیار فشرده تر هستند. از طرف دیگر، تنظیم قوانین برای یک رمز گذاری غیرمستقیم می تواند منجر به یک سو گیری شدید در فضای جستجو شود. بنابراین، ایجاد یک رمز گذاری غیرمستقیم بدون دانش اساسی در مورد نحوه استفاده از رمز گذاری بسیار دشوارتر است.

به دلیل سادگی و طبیعی تر شدن شرایط مسئله، الگوریتم نیت از روش مستقیم استفاده می کند تا در عین پیچیدگی بتوان درک شفافی از آن داشت. در شکل زیر ساختار ژنوتیپ ها و فنوتیپ ها مشاهده می شود:

¹⁶ Deep Neuroevolution

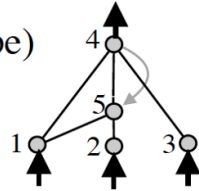
¹⁷ Node

¹⁸ Genotype

¹⁹ Phenotype

Genome (Genotype)							
Node	Node 1	Node 2	Node 3	Node 4	Node 5		
Genes	Sensor	Sensor	Sensor	Output	Hidden		
Connect.	In 1	In 2	In 3	In 2	In 5	In 1	In 4
Genes	Out 4	Out 4	Out 4	Out 5	Out 4	Out 5	Out 5
	Weight 0.7	Weight-0.5	Weight 0.5	Weight 0.2	Weight 0.4	Weight 0.6	Weight 0.6
	Enabled	DISABLED	Enabled	Enabled	Enabled	Enabled	Enabled
	Innov 1	Innov 2	Innov 3	Innov 4	Innov 5	Innov 6	Innov 11

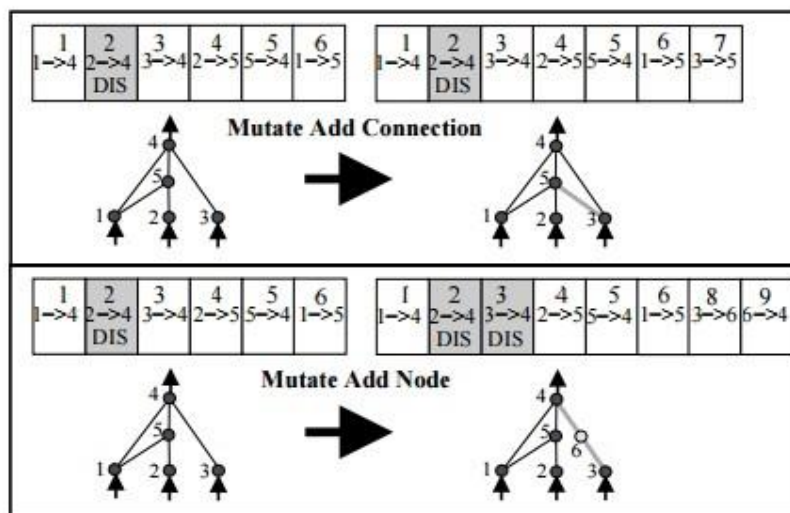
Network (Phenotype)



شکل ۷- ساختار ژنوتیپ ها و فنوتیپ ها

۳-۴ جهش

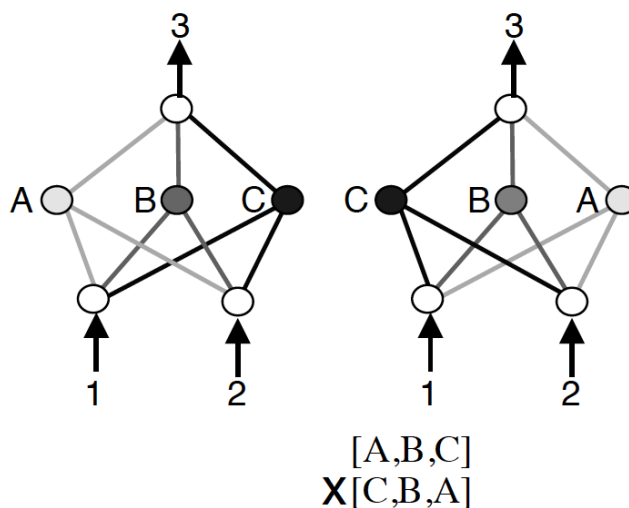
در این الگوریتم، جهش می تواند اتصالات موجود را تغییر دهد یا ساختار جدیدی را به شبکه اضافه و حتی تغییر دهد. در صورتی که یک اتصال جدید بین گره شروع و پایان یا به عبارتی اول یا آخر رشته ژن ها اضافه شود، به طور تصادفی وزنی به آن اختصاص داده می شود. همچنین اگر یک گره جدید در جایی وسط رشته اضافه شود، بین دو گره که قبلاً متصل شده اند قرار می گیرد و اتصال قبلی غیرفعال می شود (اگرچه هنوز در آن ژن وجود خارجی دارد). گره شروع قبلی با وزن اتصال قدیمی به گره جدید و گره جدید با وزن ۱ به گره انتهایی قبلی مرتبط می شود.



شکل ۸- روش های مختلف برای جهش در نقاط مختلف رشته ژن ها

۴-۳ معاونت های رقابتی^{۲۰}

از مشکلات الگوریتم ژنتیک تعریف تابع ترکیب است به گونه ای که ایده آن است که نه تنها ترکیب کورکورانه لزوماً منجر به پیشرفت شبکه نشده، بلکه می تواند به مرور زمان شبکه ای غیر قابل اعتماد و غیر کاربردی تشکیل دهد. همانگونه که در شکل زیر قابل مشاهده است، اگر دو شبکه به گره های مرکزی وابسته باشند که هر دو از شبکه دوباره ترکیب شوند، مشکل بزرگی داریم که منجر به از دست رفتن اطلاعات می شود و مقاله های الگوریتم نیت آنرا "معاونت های رقابتی" می نامند [6]:



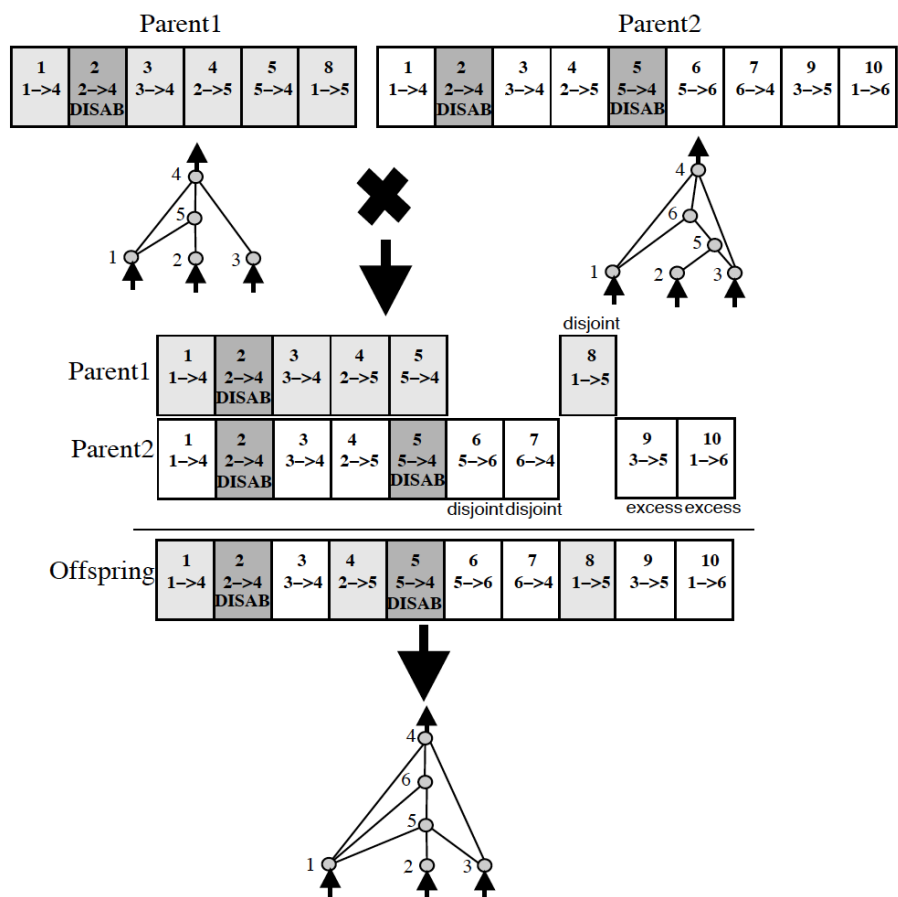
Crossovers: $[A,B,A]$ $[C,B,C]$
(both are missing information)

شکل ۹- از دست رفتن اطلاعات در دو شبکه که به گره مرکزی وابسته هستند

از طرفی، ژن ها می توانند اندازه های مختلفی داشته باشند. چگونه ژن هایی را که از نظر اندازه با هم سازگار نیستند را تراز کنیم؟ در زیست شناسی، از طریق ایده ای به نام همسانی به این امر رسیدگی می شود. همسانی، همراستا سازی کروموزومها بر اساس ژن های مطابق برای یک صفت خاص است. هنگامی که این اتفاق می افتد، تابع ترکیب می تواند با احتمال خطای بسیار کمتری نسبت به زمانی که کروموزومها کورکورانه با هم مخلوط شده باشند، اتفاق بیفتد.

الگوریتم نیت با استفاده از نشانه های تاریخی با این مشکل مقابله می کند. با علامت گذاری تکامل های جدید با یک عدد تاریخی، در هنگام ترکیب دو والد، می توان این کار را با شانس بسیار کمتری برای ایجاد فرزند های غیر عملیاتی انجام داد. هر ژن را می توان همراستا و (به طور بالقوه) ترکیب کرد. هر بار که یک گره جدید یا نوع جدیدی از اتصال رخ می دهد، یک علامت گذاری تاریخی اختصاص داده می شود که اجازه می دهد در هنگام پرورش دو والد همسویی آسان تری داشته باشیم. شکل زیر توصیف بهتری از این موضوع دارد [6]:

²⁰ Competing Conventions



شکل ۱۰- استفاده از نشانه های تاریخی برای همسو سازی آسان تر والد ها

۳-۴ گونه سازی^{۲۱}

ایده جالبی در توسعه الگوریتم نیت مطرح شد که بیشتر تحولات جدید، موارد خوبی نیستند. در واقع، افزودن یک فرزند جدید - که می تواند گره یا اتصال جدیدی باشد - قبل از بهینه سازی ژن ها، اغلب منجر به عملکرد ضعیف تر فرزند می شود. این امر ساختارهای جدید را در معرض بحران از بین رفتن قرار می دهد و نیازمند روشی هستیم تا از ساختار های تازه شکل گرفته (و در حال شکل گرفتن) محافظت کنیم تا به آنها اجازه دهیم قبل از حذف کامل آنها از جمعیت، بهینه سازی شوند. نیت مفهوم گونه سازی را پیشنهاد می کند.

گونه سازی به سادگی جمعیت را بر اساس شباهت ژن ها و اتصالات به چندین گونه تقسیم می کند. عملکردی برای تصمیم گیری در مورد چگونگی گونه سازی در مقاله های معتبر ارائه شده است، اما بخش مهمی که باید به آن توجه کرد آن است که افراد در یک جمعیت فقط باید با افراد دیگر در همان گونه مشابه رقابت کنند. این مفهوم ایجاد و بهینه سازی ساختار جدید را با خطر کمتری برای حذف شدن پیش از توسعه و بهینه سازی فراهم می کند.

²¹ Speciation

فصل چهارم

استفاده از الگوریتم نیت در حل مسائل

۴- استفاده از الگوریتم نیت در حل مسائل

در فصل دوم با استراتژی های تکاملی و سپس با الگوریتم های موجود نظیر الگوریتم ژنتیک و جست و جو تازگی آشنا شدیم. در فصل سوم با الگوریتم تکامل عصبی عمیق آشنا و ویژگی های انقلابی که آنرا متمایز می کند آشنا شدیم. حال می توان با استفاده از آموخته های این دو فصل مسئله های مختلفی را حل کنیم.

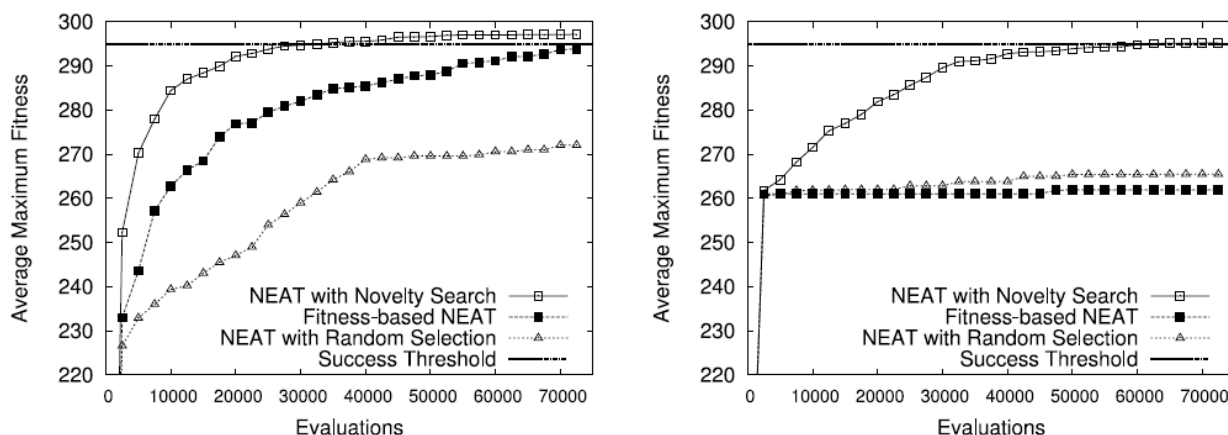
آزمایش های گوناگونی با استفاده از ترکیب های مختلف الگوریتم تکامل عصبی عمیق در مسائل واقعی برای مقایسه بهترین ساختار برای این الگوریتم و انتخاب بهترین ساختار طراحی و انجام شده است. در آنها بخصوص از الگوریتم های زیر استفاده شده است:

- الگوریتم نیت با استفاده از جست و جو تازگی
- الگوریتم نیت با استفاده از الگوریتم ژنتیک
- الگوریتم نیت با استفاده از انتخاب تصادفی

که در ادامه به چندین آزمایش بارز انجام شده توسط این الگوریتم ها می پردازیم:

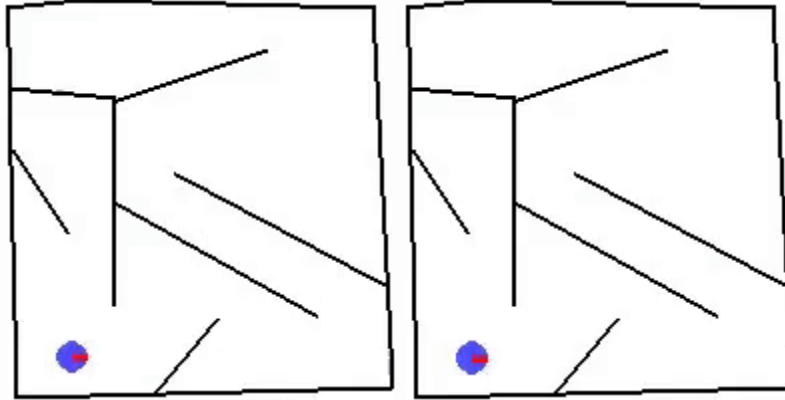
۴-۱ شبکه ماریچ فرینده

این آزمایش بر روی ترکیب های مختلف الگوریتم تکامل عصبی عمیق در دو نقشه مختلف انجام شد:



شکل ۱۱- مقایسه الگوریتم ها در حل مسئله ماریچ

مطابق انتظار برای مسائلی از این قبیل استفاده از مفهوم جست و جو تازگی در الگوریتم تکامل عصبی عمیق تاثیر بسزایی در یافتن مسیر دارد که باعث شده در کمترین زمان (تعداد نسل) مسئله حل شود و سخت شدن نقشه آزمایش تاثیر زیادی بر روی عملکرد این الگوریتم ندارد. اما برای الگوریتم های دیگر بزرگ شدن نقشه به طور قابل توجهی عملکرد الگوریتم های دیگر را کاهش می دهد به طوری که دیگر قادر به رسیدن به مقصد نیستند [7]:

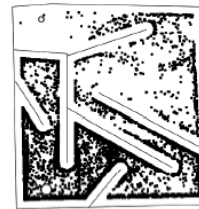


شکل ۱۲- مقایسه دو الگوریتم در شبیه سازی مارپیچ. در سمت راست الگوریتم با تابع هدف و در سمت چپ با جستجو تازگی قرار دارند [8]

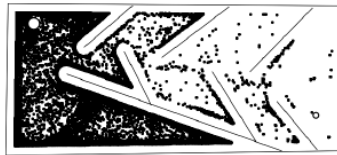
این آزمایش ۴۰ بار صورت گرفت و الگوریتم نیت با جست و جو تازگی در ۳۹ بار توانست به مقصد برسد، در صورتی که الگوریتم های دیگر کمتر از ۱۰٪ مواقع موفق به حل مسئله و یافتن مسیر شدند. شبیه سازی مسیر های طی شده در الگوریتم ها را مشاهده می کنید:



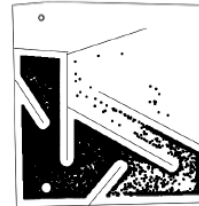
(a) Medium Map Novelty



(b) Hard Map Novelty



(c) Medium Map Fitness



(d) Hard Map Fitness

شکل ۱۳- شکل های مارپیچ بعد از ۲۵۰ هزار نسل

۲-۴ آزمایش دویا انسان نما

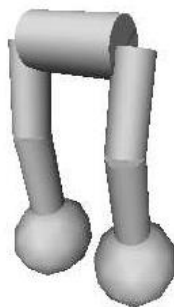
در حالی که دامنه مسئله مارپیچ نشان می دهد که جستجو برای تازگی با آزمودن مسیر های جدید می تواند سریع تر به مقصد برسد، یک سوال منطقی این است که آیا جستجوی جدید صرفا برای دامنه مسئله پیشین مناسب است و یا آیا می توان آن را با موفقیت در حوزه های دشوارتر آزمایش کرد؟

بنابراین نیاز به آزمایش جستجوی جدید بر روی یک مشکل شناخته شده با دشواری قابل توجه نسبت به مسئله مارپیچ داریم. شهود پشت چنین تلاشی این است که روش جستجوی جدید ممکن است در چنین حوزه ای نیز موفق باشد زیرا از آنجایی که

فصل چهارم – استفاده از الگوریتم نیت در حل مسائل

جستجوی جدید هدف را نادیده گرفته و به ساختارهای خلاقانه بهای بیشتری می دهد، ممکن است بتواند در حوزه های دشوار به سادگی با جستجوی تازگی موفق شود.

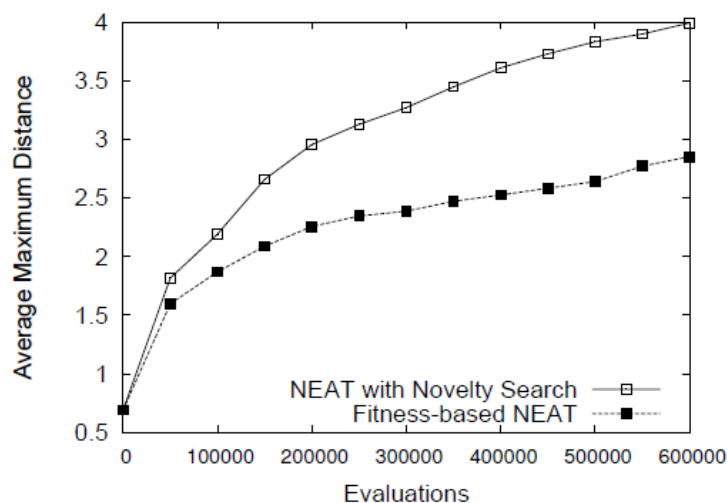
آزمایش دوپا به صورت زیر عمل می کند. یک ربات دوپا برای مدت زمان ثابتی (که در این شبیه سازی ۱۵ ثانیه در نظر گرفته شده) توسط یک شبکه عصبی کنترل می شود و سپس موفقیت شبکه برای ایستاده نگه داشتن دوپا ارزیابی می شود. از طرفی هدف آزمایش آن است که ربات بیشترین فاصله ممکن را از محل شروع بدون افتادن حرکت کند.



شکل ۱۴- یک دوپا که توسط یک شبکه عصبی قادر به ایستادن است

شبکه عصبی که دوپا را کنترل می کند تنها دو ورودی دارد که حسگرهای تماسی هستند که به هر پا سیگنال می دهند که آیا با زمین تماس دارد یا خیر. پراکندگی ورودی مشکل را دشوارتر می کند زیرا شبکه هیچ اطلاعاتی در مورد جهت ربات یا زوایای جاری اتصالات آن ندارد [9].

بعد از ۵۰ تست متوالی شبکه دوپا با استفاده از مفهوم جست و جو تازگی به طور میانگین در هر ۱۵ ثانیه حدوداً ۱۳.۷ متر حرکت می کند؛ این در حالی است با استفاده از تابع هدف حدوداً ۶.۷ متر حرکت داشتیم. نتایج در شکل زیر قابل مشاهده است:



شکل ۱۵- مقایسه عملکرد هر دو الگوریتم در ۵۰ تست متوالی

۳-۴ آزمایش بازی آتاری

آموزش شبکه های عصبی عمیق برای بازی آتاری با نگاشت مستقیم پیکسل ها برای تولید اقدامات یک شاهکار مشهور بود که مسلماً عصر یادگیری تقویتی عمیق را آغاز کرد و درک ما را از دشواری دامنه های یادگیری عمیق که یادگیری ماشین می تواند با آن مقابله کنند گسترش داد.

در این بخش چگونگی عملکرد یک شبکه تودرتو عمیق توسط یک الگوریتم ژنتیک بسیار ساده در مقایسه با شبکه تودرتو عمیق های آموزش دیده توسط خانواده های اصلی استراتژی های تکاملی را بررسی می کنیم و نتیجه می گیریم که همیشه باز گذاشتن دست الگوریتم برای یافتن ساختار های جدید و خلاقانه لزوماً استراتژی خوبی برای حل مسئله نیست و گاهی نیاز داریم تا تابع هدف پرنرنگی را طراحی کنیم.

به دلیل محدودیت منابع محاسباتی، ما نتایج ۱۳ بازی آتاری را با هم مقایسه کرده و نتایج را مشاهده می کنیم:

	DQN	ES	A3C	RS	GA	GA
Frames	200M	1B	1B	1B	1B	6B
Time	~7-10d	~ 1h	~ 4d	~ 1h or 4h	~ 1h or 4h	~ 6h or 24h
Forward Passes	450M	250M	250M	250M	250M	1.5B
Backward Passes	400M	0	250M	0	0	0
Operations	1.25B U	250M U	1B U	250M U	250M U	1.5B U
amidar	978	112	264	143	263	377
assault	4,280	1,674	5,475	649	714	814
asterix	4,359	1,440	22,140	1,197	1,850	2,255
asteroids	1,365	1,562	4,475	1,307	1,661	2,700
atlantis	279,987	1,267,410	911,091	26,371	76,273	129,167
enduro	729	95	-82	36	60	80
frostbite	797	370	191	1,164	4,536	6,220
gravitar	473	805	304	431	476	764
kangaroo	7,259	11,200	94	1,099	3,790	11,254
seaquest	5,861	1,390	2,355	503	798	850
skiing	-13,062	-15,443	-10,911	-7,679	†-6,502	†-5,541
venture	163	760	23	488	969	†1,422
zaxxon	5,363	6,380	24,622	2,538	6,180	7,864

شکل ۱۶ – مقایسه امتیاز ثبت شده توسط الگوریتم های مختلف یادگیری عمیق

مقایسه عملکرد بین الگوریتم ها ذاتاً چالش برانگیز است، اما سعی می کنیم با نشان دادن تخمین هایی برای مقدار بخش های مختلف آنها را ارزیابی کنیم. همانگونه که پیدا است هر الگوریتم بسته به طراحی و شیوه حل مسئله، در مسئله خاصی عملکرد قابل توجهی از خود نشان می دهد و نمی توان الگوریتم ثابتی برای حل تمامی مسائل طراحی کرد.

فصل پنجم

نتیجه گیری و پیشنهاد ها

۵- نتیجه گیری و پیشنهاد ها

در بخش پایانی گزارش، جمع بندی و مروری بر سیر مطالب عنوان شده در گزارش خواهیم داشت. همچنین نتایج حاصل را بیان کرده و پیشنهاد هایی در راستای ادامه کار ارائه می دهیم.

۵-۱ نتیجه گیری

در این گزارش با هدف تمرکز بر روی ارائه روش نوینی برای آموزش شبکه های عصبی که بیشتر از روش های سنتی بر روی طبیعی بودن تکامل تاکید داشته، در فصل اول سعی کردیم مفاهیم استراتژی های تکاملی را بازگو کنیم. سپس با مطالعه چندین ویژگی بارز از جمله موازی سازی، مقیاس پذیری و پارامتر سازی بر انعطاف پذیری این استراتژی ها تاکید کردیم. در نهایت با معرفی دو الگوریتم ژنتیک و جست و جو تازگی به جمع بندی خوبی درباره این استراتژی حل مسئله رسیدیم.

در فصل دوم با معرفی الگوریتم نیت به عنوان الگوریتم جایگزین روش سنتی کاهش نزولی شیب گرادین برای آموزش شبکه های عصبی و بیان ویژگی های مهم آن از جمله نحوه رمز گذاری آن، نحوه ایجاد جهش و اعمال معاونت های رقابتی و در نهایت گونه سازی سعی بر گسترش این موضوع، بیان قدرت و انعطاف پذیری این الگوریتم داشتیم.

در فصل سوم با نشان دادن چندین آزمایش و شبیه سازی های صورت گرفته در این زمینه مانند مسئله ماریچ فریبده و شبیه سازی دوبا به اهمیت استفاده از الگوریتم های جست و جو تازگی و الگوریتم هایی که انعطاف پذیری بالایی برای یادگیری روند های مختلف داشته پرداخته و سعی بر کم تر کردن نقش انسان در تنظیم کردن پارامتر ها و پر رنگ کردن نقش خود مدل برای انجام این کار داشتیم.

مهم ترین درسی که از این گزارش و بررسی شیوه های مختلف حل مسئله می گیریم نقش خلاقیت در خلق مدل های هوشمند تر است و اجبار به کمتر کردن نقش انسان در کنترل این مدل ها است. به گونه ای که مدل ها قابلیت وقف دادن خود با شرایط یادگیری حالت های مختلف را داشته باشند و بدین صورت به هدف این گزارش رسیده که الگوریتم های مختلف با توجه به شرایط مسئله و عوامل فراوان دیگر می توانند خروجی متفاوتی بر روی مسئله داشته باشند؛ به گونه ای که الگوریتم جست و جو طبیعی به دنبال روش های خلاقانه است در صورتی که اینگونه روش ها در برخی مسائل باعث هدر رفت منابع می شود؛ پس استفاده از الگوریتم ژنتیک به صرفه تر است. از طرفی در برخی مسائل نیازمند خلاقیت بالا برای حل مسئله هستیم که الگوریتم ژنتیک را ناکارآمد می سازد.

۵-۲ پیشنهاد ها

روش های متعددی برای آموزش شبکه های عصبی وجود دارند که نه تنها سریع تر و بهینه تر هستند بلکه انعطاف پذیری بیشتری دارند. ما در این مقاله چندین الگوریتم را مورد مطالعه قرار داده و در نهایت ترکیب کردیم اما نباید فراموش کرد که این حوزه پژوهش بسیار فعال و امیدوار کننده است. از طرفی همچنان چالش های متعددی وجود دارند که الگوریتم های مطرح شده در این گزارش لزوماً کارآمد ترین روش برای مواجهه با آنها نیستند.

فصل پنجم - نتیجه گیری و پیشنهاد ها

یکی از مهم ترین دستاورد های چند سال اخیر، استفاده از روش ترکیبی^{۲۲} برای ترکیب مدل ها و خروجی آنها است. در این روش که زیر شاخه های زیادی دارد مدل ها (و گاه اطلاعات) به شیوه های گوناگونی با یکدیگر ترکیب شده تا خروجی بهتری دهد. طبق مقالات معتبر، استفاده از این روش به تنهایی میتواند دقت مدل های یادگیری ماشین را بین ۵-۱۰ درصد افزایش دهد. استفاده از این روش برای آموزش شبکه های عصبی عمیق حوزه ای نوپا بوده و متخصصین همچنان در حال جمع آموری اطلاعات و آزمایش مدل های مختلف هستند.

²² Ensemble

- [1] T. Salimans, J. Ho, X. Chen, S. Sidor, and I. Sutskever, "Evolution Strategies as a Scalable Alternative to Reinforcement Learning," arXiv, arXiv:1703.03864, Sep. 2017. doi: 10.48550/arXiv.1703.03864.
- [2] E. Conti, V. Madhavan, F. P. Such, J. Lehman, K. O. Stanley, and J. Clune, "Improving Exploration in Evolution Strategies for Deep Reinforcement Learning via a Population of Novelty-Seeking Agents," arXiv, arXiv:1712.06560, Oct. 2018. doi: 10.48550/arXiv.1712.06560.
- [3] X. Zhang, J. Clune, and K. O. Stanley, "On the Relationship Between the OpenAI Evolution Strategy and Stochastic Gradient Descent," arXiv, arXiv:1712.06564, Dec. 2017. doi: 10.48550/arXiv.1712.06564.
- [4] F. P. Such, V. Madhavan, E. Conti, J. Lehman, K. O. Stanley, and J. Clune, "Deep Neuroevolution: Genetic Algorithms Are a Competitive Alternative for Training Deep Neural Networks for Reinforcement Learning," arXiv, arXiv:1712.06567, Apr. 2018. doi: 10.48550/arXiv.1712.06567.
- [5] H. Mania, A. Guy, and B. Recht, "Simple random search provides a competitive approach to reinforcement learning," arXiv, arXiv:1803.07055, Mar. 2018. doi: 10.48550/arXiv.1803.07055.
- [6] "NEAT: An Awesome Approach to NeuroEvolution | by Hunter Heidenreich | Towards Data Science." <https://towardsdatascience.com/neat-an-awesome-approach-to-neuroevolution-3eca5cc7930f> (accessed Jun. 03, 2022).
- [7] J. Lehman, J. Chen, J. Clune, and K. O. Stanley, "Safe Mutations for Deep and Recurrent Neural Networks through Output Gradients," arXiv, arXiv:1712.06563, May 2018. doi: 10.48550/arXiv.1712.06563.
- [8] K. O. Stanley, "Welcoming the Era of Deep Neuroevolution," *Uber Engineering Blog*, Dec. 18, 2017. <https://eng.uber.com/deep-neuroevolution/> (accessed Jun. 03, 2022).
- [9] N. Narayanan, "On Genetic Algorithms: Why Novelty Search is important," *Medium*, Oct. 04, 2020. <https://medium.com/@nniranjhana/on-genetic-algorithms-why-novelty-search-is-important-6d2879d3ed81> (accessed Jun. 03, 2022).