

تمرین اول هوش محاسباتی

(۱)

پایه تکامل: ۱- تنوع ۲- انتخاب طبیعی

تنوع در مقابل انتخاب قرار دارد. انتخاب تنوع را کنترل میکند و تنوع جلوی انتخاب را میگیرد.

تنوع، شرط تکامل است.

این دو عامل با هم ترید آف دارند و هر دو برای تکامل ضروری می باشند.

هدف تکامل بقای ژن های قابل انتقال است.

انتخاب یعنی: چه انتخاب هایی به تکامل می انجامند؟

(۲)

دانش، دانش محلی است. و اگر جمعیت اولیه بایاس داشته باشد، جواب ما در می نی مم لوکال گیر می افتد.

به همین دلیل، باید حتما در انتخاب خود، توزیع نرمال را در نظر بگیریم.

(۳)

دو حالت خاص وجود دارد:

- در Generational یعنی میو و لاندا با هم برابرند.
- در حالت steady state لاندا برابر با یک هست و از روش میو + لاندا استفاده می شود.

(۴)

شباهت:

- هر دو هزینه ی سورت دارند: $o(n)$
- هر دو رو، مبتنی بر شایستگی می باشند.

روش رولت ویل:

- انتخاب یک عدد بین 0,1
- در عمل در تعداد کم در رولت ویل پیشرفتی نداریم.

روش SUS:

- از لحاظ ریاضی، معادل رولت ویل است. اما ترتیب موجودات تصادفی خواهد بود.

- انتخاب یک عدد بین 0 و $1/n^2$
- هر N^2 موجود، یکبار انتخاب میشود.
- برای هر موجود یک احتمال تعریف میکنیم و برای هر موجود به اندازه $1/n^2$ آن احتمال، روی خطکش فضا میدهم.

(۵)

وقتی از $u, landa$ بهره میگیریم، تنوع بالا می رود و از می نیمم محلی فرار میکنیم. (این دو عامل یعنی فراموشکاری الگوریتم) تنوع بالا برود: همگرایی کند میشود.

(۶)

ES، برعکس روش GA که به بازترکیبی اهمیت بیشتری میداد، اول به جهش اهمیت میدهد بعد به بازترکیبی. مثلا اگر دووالد دقیقا مثل هم باشند، بازترکیبی برای آن ها تنوع را کم میکند. اما اگر اول جهش صورت بگیرد، دیگر هیچ دو والدی مثل هم نخواهند بود. پس تضمین میشود که:

$$PS < 0.4$$

در واقع ES بر این تاکید دارد که اول جهش انجام شود، سپس اندازه قدم ها تنظیم شود. تنظیم اندازه ی قدم ها توسط یک روند تکاملی انجام میگردد. (self adaptive یا خودتطبیقی). در این رویکرد، اندازه ی قدم باید به طور تدریجی، تنظیم شود. میدانیم:

$$X_i = X_i + \sigma \cdot N(0,1)$$

اگر (سیگما) σ را همینجور دلخواه تغییر دهیم و بالا پایین کنیم، متوجه نمیشویم چه مقداری از آن بهینه است و کدام مقدار به درد ما نمیخورد.

اگر σ اندازه قدم ها باشد، میتوانیم برای تغییرات و جایگزینی اندازه قدم (آپدیت کردن)، یک فرمول بسازیم:

$$\sigma^* = \sigma \cdot e^{-T \cdot N(0,1)}$$

در این فرمول، اگر T خیلی کوچک باشد یادگیری کندتر خواهد بود و مقدار عبارت نمایی تقریبا برابر با یک ۱ است. و اگر خیلی بزرگ باشد، یادگیری سریع است.

حالا سوال این است که چگونه میفهمیم یک مقدار σ مقدار مناسبی است یا خیر؟ جواب: از روی X . (راه دیگری هم نداریم).

اگر قرار باشد σ دائم تغییر کند و آپدیت شود، ما نمیتوانیم بفهمیم که کدام مقدار مناسب بوده است، چون σ سریع آپدیت شده است.

پس علاوه بر اینکه تغییر دادن σ باید تدریجی انجام شود، امکان اینکه اول X جهش پیدا کند هم وجود ندارد و حتما باید اول σ جهش یابد. چرا که تنها در این حالت است که ما میتوانیم از روی مقدار X ، بفهمیم مقدار بهینه ی σ چیست و سپس بر اساس این مقدار، X متناظر با آن را جهش دهیم.

به همین علت ابتدا سیگما و سپس X را جهش میدهم.

"پایان"