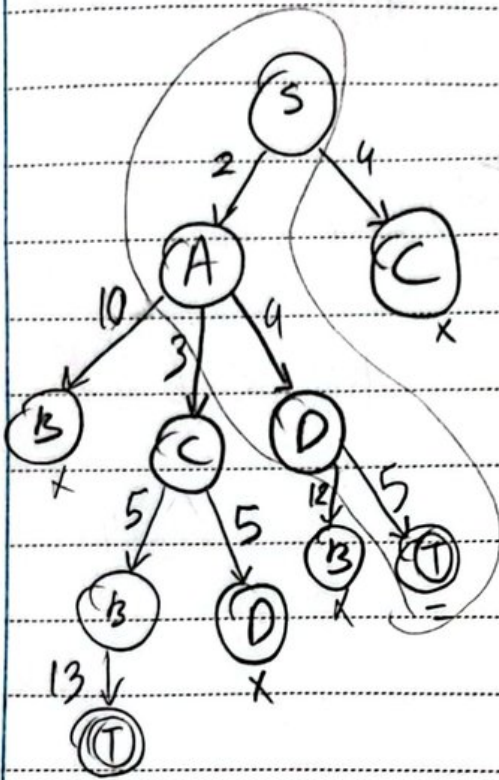


تمرین اول آلوری

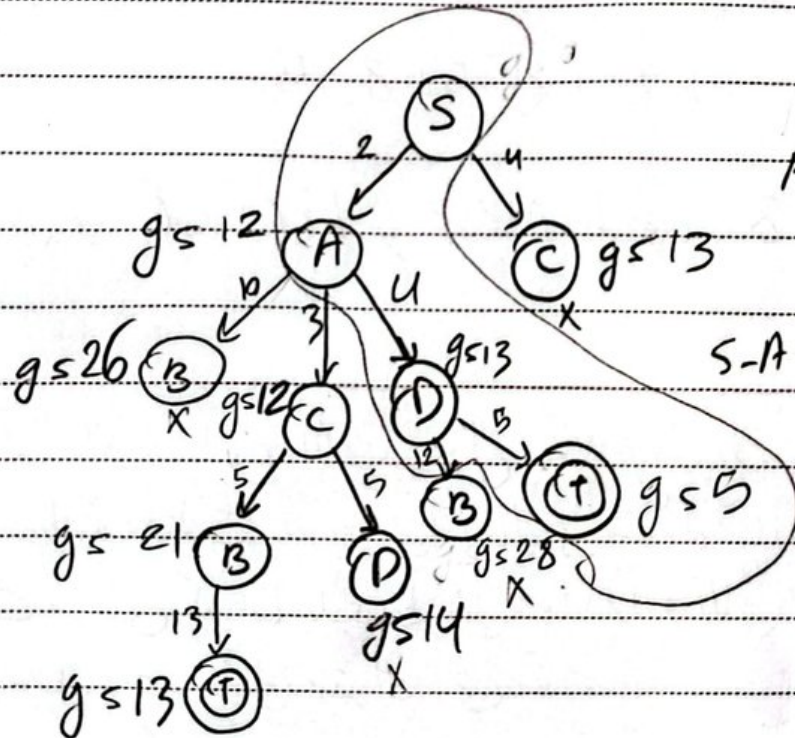
UCS



الف)

S-A-D-T
cost: 5

A*

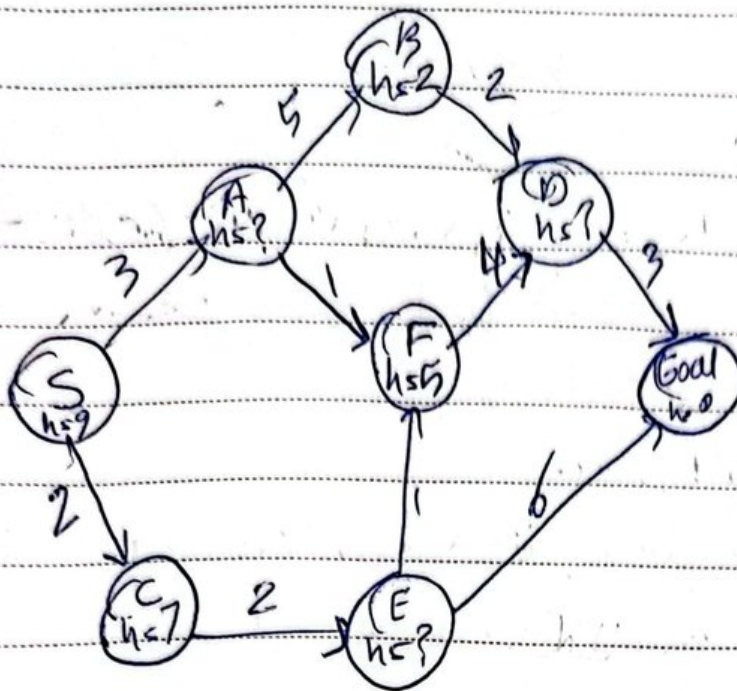


ب)

S-A-D-T
cost: 5

ج) به تبع مسائل است ، به این دلیل که تابع هزینه یکسان است .

2- با یار و قابل قبول



قابل قبول

$$h(A) < 8 \text{ یا } 10$$

$$h(D) < 3$$

$$h(E) < 6 \text{ یا } 8$$

با یاری:

$$h(A) < h(F) + 7 \leq 12$$

$$h(D) < 3$$

$$h(E) < 5 + 8 \leq 13$$

استدلال بالا یاری: $h(E) < 6$ و $h(D) < 3$ و $h(A) < 8$

3- الف) اگر ایما این الگوریتم معادل الگوریتم Hill climbing است. چون در ایما فقط یک حالت در هر لحظه دنبال می شود که مشابه Hill climbing است. جایی که تنها یک نقطه فعلی وجود دارد و تلاش برای بهبود در صفت یا جابجایی همسایگان بهتر انجام می شود.

(ب) اگر $k=0$ ، این الگوریتم به BFS تبدیل می شود.

در واقع با این مقدار، ما به طور همزمان همه حالات ممکن را در نظر می گیریم و بهترین را انتخاب می کنیم که اساساً جستجوی تمام فضای حالت است.

(ج) اگر $T=0$ ، این الگوریتم مشابه Hill climbing است چون هیچ احتمالی برای انتخاب حرکاتی که منجر به وضعیت بدتر می شوند وجود ندارد و فقط به سمت همایه هایی که بهبود وضعیت دارند حرکت داریم.

(د) اگر ما با سرعت کاهشی (همی) ممکن است الگوریتم قبل از اینکه فرصت کافی برای جستجوی مناسب فضای حالت را داشته باشد، در یک ماکزیمم محلی گیر کند و به یک راه حل غیر بهینه محدود شود.

اگر ما در یک مقدار مثبت ثابت نگه داریم، الگوریتم تمایل بیشتری به پذیرش حرکات بدتر در طول جستجو دارد که می تواند به افزایش فضای حالت و احتمال پیدا کردن راه بهینه تر ختم شود.

اما اگر ما زیاد بایستیم، الگوریتم به یک جستجوی تصادفی بدون هدف تبدیل می گردد.

۴- الف) معنی train کاهش خطای دسته آموزشی را نشان می دهد
زیرا مدل با گذشت زمان یاد می گیرد.

آرمر در دسته validation از توضیح یکسانی آمده باشند و validation خطای بیشتری داشته باشد. این موضوع نشانه overfitting است. برای حل این مشکل، می شود از تکنیک هایی مانند regularization، drop out در شبکه های عصبی یا به طور ساده اضافه کردن دیتاهای آموزشی متنوع تر، استفاده کرده.

ب) در این صورت، خروجی های این دو دسته تفاوت در خطا را نشان می دهد. (difference in loss). و نشان دهنده این است که مدل ما توانایی تعمیم داده های مختلف را ندارد.

راه حل این مشکل نیز این است که مدل توسط داده های متفاوت و نمونه هایی از هر دو توزیع، آموزش ببیند. استفاده از پارامتر ارزیابی متفاوت نیز می تواند موثر باشد.

$$A = ABCDEFGH$$

-5

$$f(x) = A + (2 \times B) + (3 \times C) - D - E - (2 \times F) + G + H$$

$$x_1 = 24513892$$

A B C D E F G H

$$f(x_1) = 2 + (2 \times 4) + (3 \times 5) - 1 - 3 - (2 \times 8) + 9 - 2 = 12$$

$$f(x_2) = 4$$

$$f(x_3) = 3$$

$$f(x_4) = 18$$

$$P(x_n) = \frac{\text{fitness}(x)}{\text{جمع fitness}}$$

احتمال انتخاب هر کروموزوم :

$$P(x_1) = 0.324$$

$$P(x_2) = 0.108$$

$$P(x_3) = 0.081$$

$$P(x_4) = 0.486$$

ب) برای x_2 و x_1 crossover انجام

$$x_1 = 24513892 \quad , \quad x_4 = 12903621$$

A B C D E F

فرزند اول : 24513621 فرزند دوم : 12903892

$$x_1 = 24513892$$

A B C D E F G H

crossover برای x_3 و x_1 :

$$x_3 = 43213205$$

A B C D E F G H

فرزند سوم : 24513892

F

فرزند چهارم : 43513205

چهارم : 12 (ج) فرزند اول : 10 درم : 20 سوم : 3

محمد یح ارزش fitness با اثراتش یافته (از 37 به 45)

(د) فرزند دوم جدید که از crossover درست آمد با مقدار 20 و بهترین fitness را دارد.

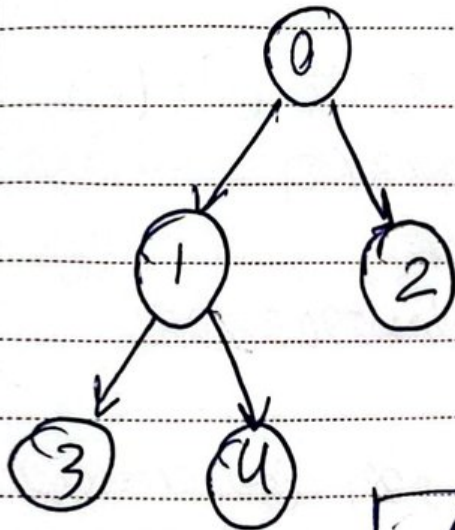
★ (ه) اگر جمعیت فعلی تنوع ژنتیکی کافی را برای رسیدن به بهترین حالت را دارا باشد، پاسخ مثبت است. در غیر اینصورت فقط با mutation امکان پذیر است. در این مسئله چون اطلاعات کافی را نداریم، مشخص نیست.

6- الف)

$$H(X) = - \sum p(x) \log_2 p(x)$$

A	B
12	0

$$\sum G(X, A) = H(X) - \sum p(x) H(X|A)$$



A	B
4	12

A	B
8	4

$$P(A) = \frac{12}{12+8} = 0.6$$

$$P(B) = \frac{8}{12+8} = 0.4$$

$$H(X) = -0.6 \log_2 0.6 - 0.4 \log_2 0.4 = 0.922$$

$$H(X|A) = -0.75 \log_2 0.75 - 0.25 \log_2 0.25 = 0.622$$

$$IG(X;A) = 0.922 - 0.622 = 0.3$$

Subject :

Date :

(ب)

1- (ب) تکنیک های اترایش رفت classification :
pruning : حذف شاخه هایی که به بهبود رفت کمک نمی کنند

2- استفاده از الگوریتم های SVM

3- اترایش تعداد نمونه ها در کلاس هایی که نمونه کمتری دارند یا کاهش تعداد نمونه ها در کلاس هایی که نمونه های بیشتری دارند

$$P(+)=\frac{1}{2} \quad P(-)=\frac{1}{2}$$

7- الف)

$$entropy = -\left(\frac{1}{2} \log_2 \frac{1}{2} - \frac{1}{2} \log_2 \frac{1}{2}\right)$$

(ب) برای ویژگی A، هر دو مقدار True و False دو بار ظاهر می شوند و در هر دو مورد توزیع کلاس یکسان است پس آنتروپی او مجموعه کلی هیچ بهره اطلاعاتی ندارد

ویژگی B، 3 بار True و 3 بار False می شود و با فرض اول 16 بهره اطلاعات برای آن 0.0817 است که کمی به عدم قطعیت کمک می کند