



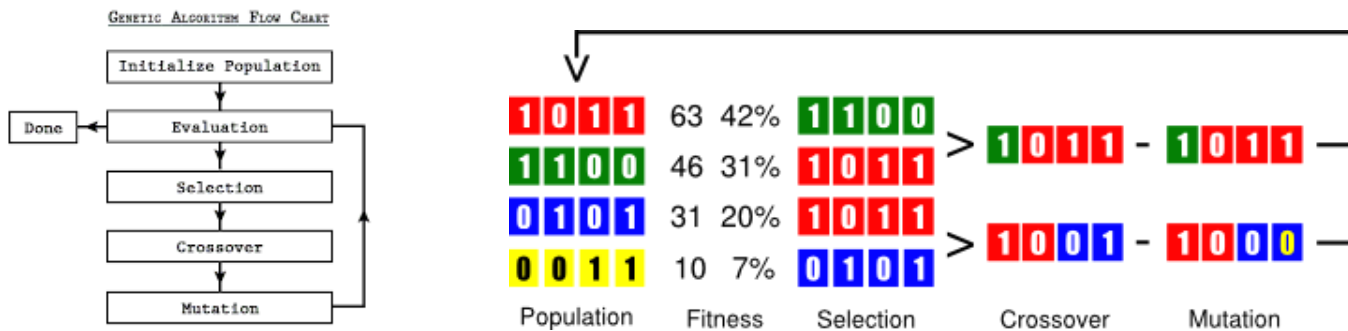
«به نام خدا»



# Artificial Intelligence

## Computer Assignment 1

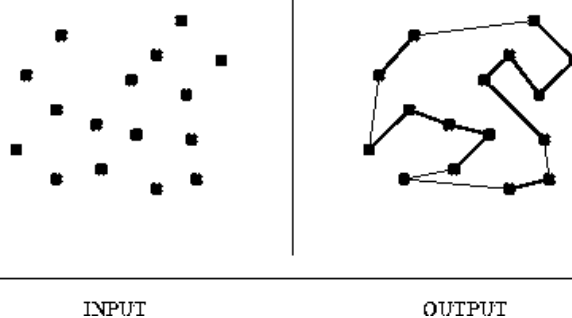
The Amazing Genetic Algorithms



این تمرین، از دو بخش تشکیل شده است. مسئله اول جنبه پوشش دادن مباحث تئوری الگوریتم ژنتیک را دارد و برای آشنایی با قسمت‌هایی مختلف آن و روش‌های مرسوم در هر قسمت تدارک دیده شده است. در مسئله دوم به صورت عملی با یکی از کاربردهای الگوریتم ژنتیک و نحوه پیاده‌سازی آن آشنا خواهید شد.

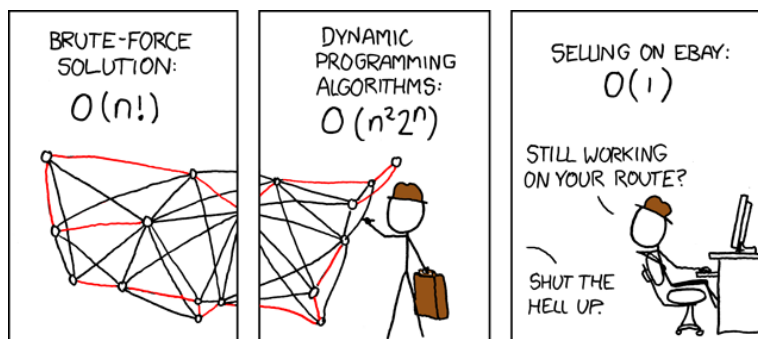
### مسئله اول) مسئله فروشنده دوره‌گرد

فروشنده‌ای برای فروش اجناسی که مدتی است خریداری نداشته، تصمیم می‌گیرد به شهرهای مختلف منطقه خود سفر کند تا مگر بتواند از این طریق برای کالاهای خود مشتری پیدا کند. از آن جایی که فرد مقتصد و صرفه‌جویی است، برای اینکه وقت و انرژی کمتری در این زمینه صرف کند، می‌خواهد قبل از شروع دوره‌گردی، ابتدا بهترین و کوتاه‌ترین مسیری که با پیمودن آن حتماً از هر شهری فقط و فقط یک بار عبور کند و در آخر به شهر خود (مکان شروع حرکت) باز گردد را بیابد.



تصویر ۱ - ورودی و خروجی مطلوب مسئله فروشنده دوره‌گرد

از شما خواسته شده راه حلی برای این مسئله (با هر تعداد شهر ورودی) ارائه دهید. بعنوان مثال یک نمونه ورودی و خروجی این مسئله در تصویر ۱ دیده می‌شود. اما متأسفانه این مسئله یک مسئله NP-hard است!



تصویر ۲- تلاش برای حل مسئله فروشنده دوره‌گرد

اینجاست که نقش الگوریتم‌های ژنتیک در مسائلی که در فضای بزرگ به دنبال راه‌حل بهینه‌ای هستند، نمایان می‌شود. البته همانطور که میدانید ممکن است الگوریتم ژنتیک به بهترین جواب نرسد اما می‌تواند تا حد خوبی به آن نزدیک شود. در این جا می‌خواهیم به کمک الگوریتم ژنتیک راه حلی برای این فروشنده پیدا کنیم.

الف) کروموزوم‌های این مسئله را چگونه تعریف می‌کنید؟ چه کروموزوم‌هایی معتبر هستند؟

ب) کروموزوم‌ها را چگونه evaluate می‌کنید؟

ج) روش‌های مختلفی از جمله Bit Flip, Scrambling و ... برای مرحله mutation وجود دارد که می‌توانید در این لینک مشاهده کنید :

[https://www.tutorialspoint.com/genetic\\_algorithms/genetic\\_algorithms\\_mutation.htm](https://www.tutorialspoint.com/genetic_algorithms/genetic_algorithms_mutation.htm)

روش mutation مورد استفاده در این مسئله چه محدودیتی دارد و کدامیک از روش‌های مطرح شده، می‌تواند برای این مسئله انتخاب مناسبی باشد؟

د) حال باید روشی برای مرحله crossover انتخاب کنیم. برای این کار نیز تکنیک‌های متعددی وجود دارد که بسته به مسئله خاصی که در حال حل کردن آن هستیم، باید روش مناسب را انتخاب کنیم.

از بین روش‌های مطرح شده در لینک زیر، کدامیک را برای حل مسئله فروشنده دوره‌گرد مناسب می‌بینید؟

<http://www.obitko.com/tutorials/genetic-algorithms/crossover-mutation.php>

## مسئله دوم) مسئله رمزگشایی رمز جایگزینی

رمزنگاری، شیوه‌ای است برای تبدیل متن خام به متن رمز شده! هرچقدر که رمزگشایی یعنی رسیدن از متن رمز شده به متن اصلی سخت‌تر باشد (یعنی در واقع هزینه زمانی بیشتری داشته باشد)، الگوریتم رمزنگاری بهتری استفاده شده است. یکی از الگوریتم‌های رمزنگاری معروف، رمزنگاری به شیوه جایگزینی است. نحوه‌ی کار این الگوریتم، همان‌طور که از اسمش بر می‌آید، اینگونه است که هر حرف در متن اصلی را با حرف دیگری جایگزین می‌کند.

کلید رمزنگاری جایگزینی، نگاشت یک به یکی از حروف الفبای انگلیسی به حروف متناظرشان در متن رمز شده است. این کلید معمولاً به صورت یک جدول نشان داده می‌شود که نمونه‌ی آن در تصویر ۳ آمده است.

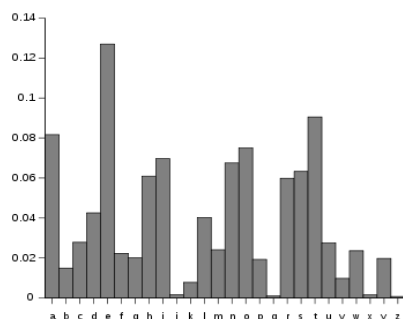
Alphabet	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Key	E	U	S	X	V	A	R	T	I	K	C	Z	M	G	Q	Y	H	O	J	P	W	D	B	N	L	F

تصویر ۳ - نمونه کلید رمزنگاری جایگزینی

هنگام رمزنگاری، هر حرف از ردیف بالا با حرف متناظرش در ردیف پایین جایگزین می‌شود و هنگام رمزگشایی، برعکس این موضوع اتفاق می‌افتد.

همانطور که احتمالاً متوجه شدید، به دلیل فضای جستجوی بزرگ مسئله، آزمون و خطا راه حل عملی برای این مسئله نخواهد بود. اینجاست که باز هم الگوریتم ژنتیک، در رویارویی با فضای جستجوی بزرگ، خودنمایی می‌کند. در گام‌های بعدی، می‌خواهیم تلاش کنیم راه حلی برای یافتن کلید رمزگشایی، به کمک الگوریتم ژنتیک بیابیم.

با کمی دقت، متوجه می‌شویم این الگوریتم رمزنگاری، آسیب‌پذیر است زیرا با وجود عوض کردن حروف متن اصلی، فرکانس استفاده حروف در متن اصلی و متن رمز شده یکی خواهد بود. به همین دلیل به کمک مقایسه کردن فرکانس تکرار حروف در متن رمز شده با جدول فرکانس حروف در زبان انگلیسی، می‌توان بخش بزرگی را از فضای جستجوی مسئله کم کرد. بعلاوه ترکیبات دو حرفی مانند *th*, *in*, *the*, *ing* و *and* نسبت به سایر ترکیب‌های دو حرفی و سه حرفی، فرکانس تکرار بیشتری دارند. (به ترکیبات دو حرفی **bigrams** و به ترکیبات سه حرفی، **trigrams** گفته می‌شود). در این جا برای سادگی فرض می‌کنیم فقط حروف (و نه علائم) رمز می‌شوند و **تنها از دانش اولیه درمورد فرکانس تکرار ترکیبات دوتایی و سه‌تایی داده شده استفاده می‌کنیم (که امتیازات مربوطه در جدول صفحه بعد آمده است).** ضمناً حروف متن اصلی فقط حروف بزرگ هستند.



in	134812613554	the	82103550112
th	133210262170	ing	43727954927
er	119214789533	and	43452082914
re	108669181717	ion	39907843075
he	106498528786	tio	32705432538
...		ent	31928292897
qy	6901470	...	
zq	6170496	jwq	10340
jx	5682177	jqy	8871
qz	4293975	zqy	8474
jq	2858953	jzq	7180
		zqq	6254

همانطور که در سوال قبل دیدید، الگوریتم ژنتیک از چند بخش تشکیل شده است. ابتدا کروموزوم‌های مسئله را تعریف کنید و سپس سعی کنید روش‌های مناسب این مسئله خاص را برای هر گام انتخاب کنید.

(در گزارشکار، باید کروموزوم و روش‌های انتخابی‌تان را به صورت مختصر توضیح دهید.)

یک نمونه فایل رمز شده که باید سعی کنید آن را رمزگشایی کنید، به همراه صورت مسئله به شما داده می‌شود.

به کمک جدول **fitness weight** زیر عملیات رمزگشایی را انجام دهید.

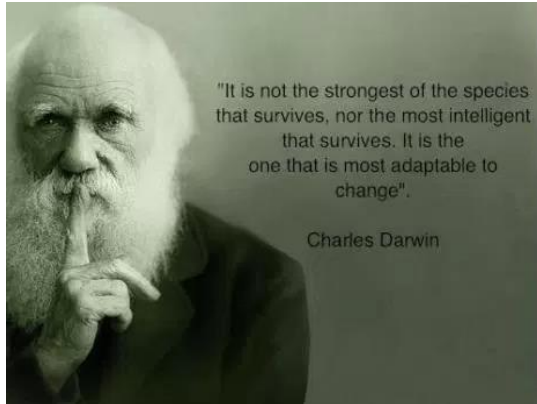
Bi/trigram	Score	Bi/trigram	Score
TH	+2	ED	+1
HE	+1	THE	+5
IN	+1	ING	+5
ER	+1	AND	+5
AN	+1	EEE	-5

توجه کنید که باید امکان تغییر آسان (یا اضافه شدن به) جدول **fitness weight** (و نیز تغییر تابع **fitness**) در کد شما فراهم باشد.

## نکات پایانی

- ۱- برای مسئله دوم میتوانید از زبان‌های C، جاوا، متلب یا پایتون استفاده کنید.
- ۲- در فایل گزارشکاری که به همراه کدهای خود آپلود میکنید باید پاسخ مسئله اول به همراه توضیحات خواسته شده در صورت سوال دوم را بنویسید.
- ۳- زمان تحویل به موقع تمرین کامپیوتری اول دوشنبه بیست و سوم اسفند می‌باشند.
- ۴- به ازای هر روز تاخیر ۱۰٪ از نمره کل تمرین کسر خواهد شد و امکان تحویل با تاخیر تا جمعه ۲۷ اسفند فراهم خواهد بود.
- ۵- سوالات خود را حتما در فروم پرسش و پاسخ درس مطرح نمایید.
- ۶- تحویل حضوری این تمرین بعد از تعطیلات عید خواهد بود. زمان آن متعاقبا اعلام خواهد شد.
- ۷- برای تحویل حضوری یکی از اعضای تیم، به صورت تصادفی، توسط دستیار آموزشی انتخاب میشود و عضو دیگر اجازه حضور در هنگام تحویل را نخواهد داشت. بنابراین هر دو عضو گروه باید به صورت کامل در جریان تمرین باشند. در غیر این صورت نمره تمرین به هیچ‌کدام داده نخواهد شد.

موفق‌تر از همیشه باشید



## کاربردهای دیگر الگوریتم ژنتیک

- در بیوانفورماتیک :
  - پیش‌بینی ساختار RNA
  - Multiple Sequence Alignment
- در الگوریتم‌های شکستن رمز
- در خوشه بندی (Clustering)
- در معماری کامپیوتر برای پیدا کردن اتصالات ضعیف
- مهندسی کنترل
- دیتاسنترها
- توپولوژی‌های شبکه کامپیوتری توزیع‌شده
- پیدا کردن مشکلات سخت‌افزاری
- پردازش و فیلتر کردن سیگنال‌ها
- کشف داروهای جدید
- کنترل کیفیت
- شبکه‌های عصبی مخصوصاً شبکه‌های بازگشتی
- بهینه‌سازی ساختار مولکولی در شیمی
- نظریه بازی‌ها
- و ...