

Makine Öğrenmesi

Doç. Dr. İlhan AYDIN

Genetik Algoritmalar

Genetik Algoritma

- 1970'li yıllarda John Holland tarafından geliştirilmiştir.
- 1989 yılında David E. Goldberg Genetik Algoritma
- Algoritma Uygulamaları üzerine klasik eser olarak kabul edilen kitabını yayınladı.
- Popülasyon tabanlı algoritmadır.
- Her birey problemin çözümü olmaya adaydır.
- Stokastik algoritmadır.
- Çözümler genellikle bit dizileri olarak kodlanır.
- Bireyleri değerlendirme fonksiyonu probleme yönelik olarak çalışan en önemli kısımdır.

Terminoloji

- Popülasyon (Topluluk): Çözüm kümesini oluşturan bireylerin tümüdür.
- Kromozom: Çözüm için kullanılan bireyler.
- Gen: Kromozom içindeki anlamlı en küçük bilgidir.
- Locus: Kromozom içindeki gen'in bulunduğu yerdir.
- Mutasyon: Bir kromozomda rastgele yapılan değişikliktir.
- Jenerasyon: Yeni bir topluluktur.
- Mating: Bireyler arasındaki eşlemedir.

İşlemler

- Çözümün veya bireyin gösterimi ve başlangıç popülasyonunun oluşturulması
- Bireylerin değerlendirilmesi (Fitness function)
- Bireylerin yeni topluluk için seçilmesi
- Yeni bir jenerasyon oluşturma

Başlangıç popülasyonunu oluşturma

- Bilinen bir çözüm kümesi alınarak başlanabilir.
- Rastgele bireyler oluşturulur.
- Belirli kriterleri sağlayan bireyler oluşturulur.
- Popülasyondaki birey sayısı genellikle 100 ile 300 arasında seçilir.
- Bireylerin Değerlendirilmesi
- Kromozomların çözüme uygunluk değerleri belirlenir
- **Bireylerin Değerlendirilmesi**
- Her problem için uygun bir değerlendirme fonksiyonu kullanılması gerekir.
- Probleme yönelik olarak çal Probleme yönelik olarak çalışan kısımdır.
- Değerlendirme her yeni jenerasyonun tüm bireyleri için yapılır.

Yeni Popülasyon Oluşturma

- Birey seçme
- Elitizm
- Seçilen iki bireyi çaprazlama
- Bir bireyin herhangi bir gen'inin mutasyon işlemiyle rastgele değiştirilmesi
- **Birey seçme**
 - Deterministik
 - Rulet tekerleği
 - Rastgele
 - Turnuva seçimi

Standart Bir Genetik Algoritma Yordamı

1. Olası çözümlerin kodlandığı bir çözüm grubu oluştur
 1. Populasyon oluşturma
2. Toplumdaki her kromozomun ne kadar iyi olduğunu bul
 1. Uygunluk fonksiyonu hesabı
3. Seçilen kromozomları eşleyerek yeniden kopyalama ve değiştirme operatörleri uygula
 1. Rulet tekerleği
 2. Turnuva seçimi
 3. Çaprazlama
 4. Mutasyon işlemleri

Standart Bir Genetik Algoritma Yordamı

4. Yeni kromozomlara yer açmak için eski kromozomlar çıkartılarak sabit büyüklükte bir toplum sağlanır
5. Tüm kromozomların uygunlukları tekrar hesaplanır
6. İşlemler tekrarlanarak verilmiş zaman içerisinde daha iyi olan yeni nesillerin oluşturulması sağlanır
7. Populasyonun hesaplanması ile en iyi bireylerin olduğu çözüm elde edilir.

Genetik Algoritma Operatörleri

- Kromozomların şifrelenmesi

- a) İkili kodlama

- Her bir kromozom 0 ve 1'lerden oluşan bit dizisidir ve ikili dizi ile ifade edilir.

Kromozom A → 10011111

Kromozom B → 11011110

- b) Permutasyon kodlama

- Bu kodlama gezgin satıcı problemi ve iş sıralama problemleri gibi sıralama problemlerinde kullanılır.

- c) Değer Kodlama

- Kromozom A → 3 5 1 2 4 7 6 0
 - Kromozom B → 5 3 4 2 1 0 6 7
 - Yolların yer aldığı problemlerde kullanılır.

Kromozom A → 3.2 6.5 0.8 9.4

Kromozom B → ABCDE

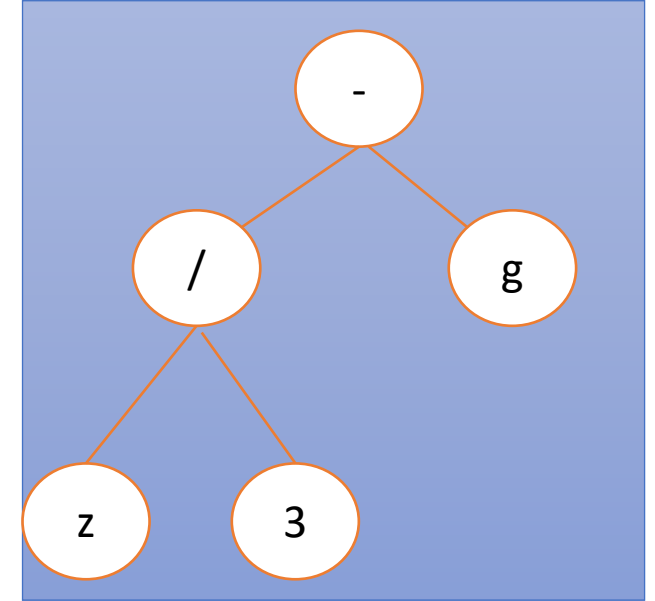
Kromozom C → (geri), (sağ), (ileri), (sol)

Genetik Algoritma Operatörleri

- **Kromozomların şifrelenmesi**

- d) **Ağaç Kodlama**

- Ağaç kodlama program geliştirmek için uygundur.
 - Örneğin Lisp ve Prolog gibi dillerde ağaç yapısı kullanılır.
 - Örneğin televizyon yarışmalarında iyi bildiğimiz verilen hedefe en yakın sonucu üretecek aritmetik ifadenin bulunmasıdır.
 - **Bir kelime Bir İşlem programı**
 - Bu ifade sayılardan ve +, -, *, / aritmetik işlemlerden oluşur.
 - İlk dört giriş sayısı 1 ile 10 arasından ve son iki sayı 25, 50, 75,100 arasından seçilir.
 - Sınırlama ise yarışmacının her sayıyı bir kez kullanabilmesidir.
 - Her zaman tam çözüm bulunmaz

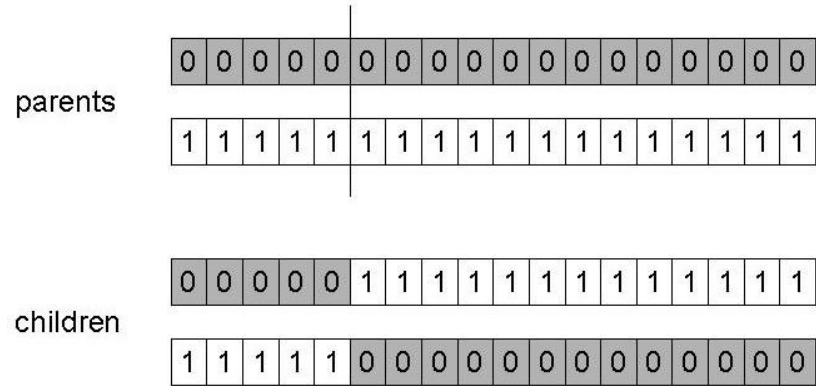


Çaprazlama ve Mutasyon

Çaprazlama

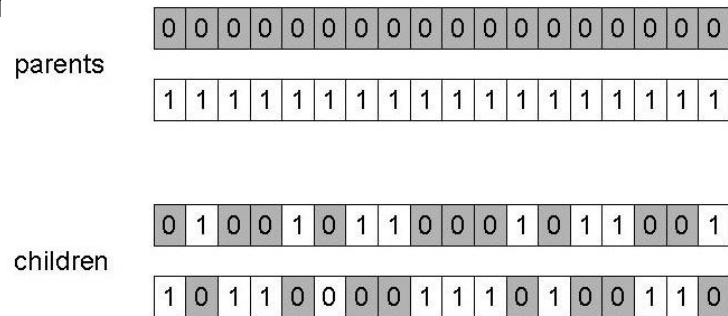
- İkili kodlanmış kromozomlarda 4 çeşit çaprazlama gerçekleştirilir.

❖ Tek noktalı çaprazlama

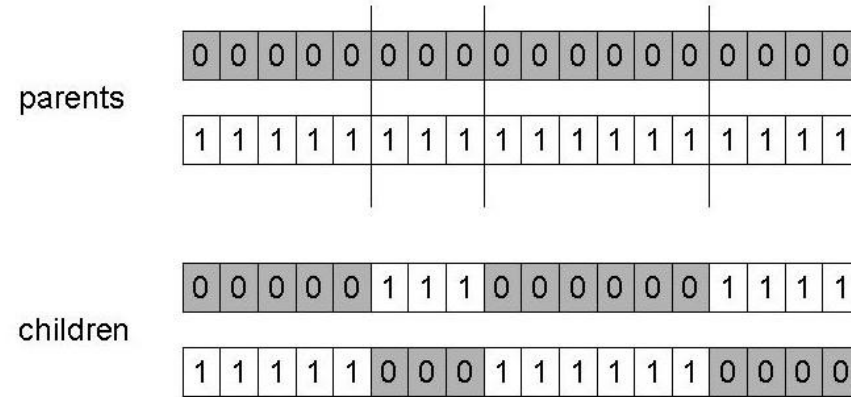


❖ Tek biçimli (uniform) çaprazlama

Bitler rastgele biçimde iki ebeveynden yeni krom



❖ Çift noktalı çaprazlama



❖ Aritmetik takas

11001011+11011111=11001001(AND) ve işlemi

Çaprazlama ve Mutasyon

- Mutasyon
 - Mutasyon işlemi, problemin popülasyondaki çözümlerinin yerel optimuma düşmesini engellemek için kullanılır
 - İkili şifreleme için, rastgele seçilen birkaç 0'dan 1'e veya 1'den 0'a değiştirilebilir.

parent

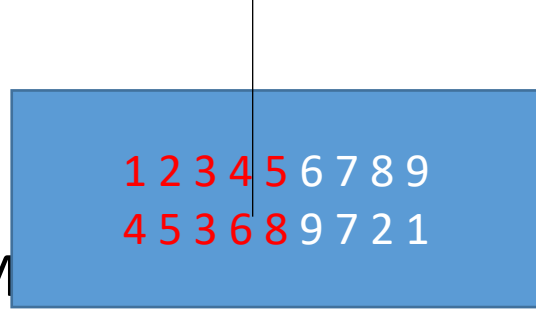
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

child

0	1	0	0	1	0	1	1	0	0	0	1	0	1	1	0	0	1
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Çaprazlama ve Mutasyon

- Permutasyon kodlanmış kromozomlarda
 - Çaprazlama: bir çaprazlama noktası seçilir. Bu noktaya kadar olan ifade birinci ebeveynden kalan ise diğer ebeveynden kopyalanarak yeni kromozom oluşturulur.



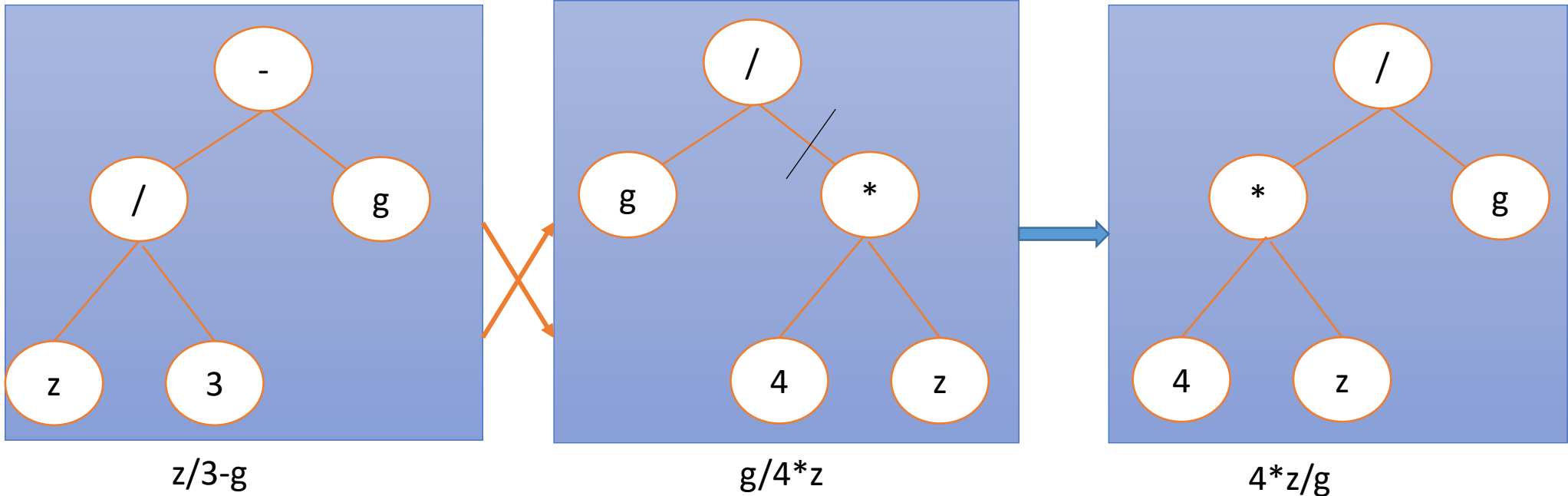
= 1 2 3 4 5 9 7 2 1

- M...ir ve bunların yerleri değiştirilir.



Çaprazlama ve Mutasyon

- Değer Kodlanmış Kromozomlarda
 - Çaprazlama: ikili kromozomlarda yapılabilen tüm takas türleri burada da yapılır.
 - Mutasyon: seçilen değerlerden küçük bir sayı çıkarılarak veya eklenerek yeni kromozom oluşturulur.
(1.29 5.68 **2.86 4.11** 5.55) → (1.29 5.68 **2.73 4.11** 5.55)
- Ağaç kodlanmış kromozomlarda
 - çaprazlama: her iki ebeveynde birer takas noktası seçilerek parçalanır. Bu noktaların altındaki kısımlar aralarında değiştirilip yeni bireyler oluşturulur.
 - Mutasyon: seçilen düğümlerdeki numaralar veya işlemler değiştirilir.



Genetik Algoritmanın Parametreleri

1. Çaprazlama olasılığı: Takasın hangi olasılıkta olduğunu belirler.
Çaprazlama olasılığı %0 ile %100 arasındadır.

Çaprazlama olasılığı %0'ın ve %100'ün anlamı nedir?

2. Mutasyon olasılığı: Mutasyonun hangi sıklıkla olacağını belirler.
Mutasyon olasılığı %0 ile %100 arasındadır.

3. Diğer parametreler:

Örnek:

- X parametresinin 0 ve 15 arasında değişken olduğu $(15x-x^2)$ fonksiyonun maksimum değerini bulalım.
- Kolaylık için biz x 'i sadece integer değerlerden alacağız. Böylece, kromozomlar dört bit'le ifade edilebilir
- Farz edelim ki; kromozom populasyon boyutu 6, çaprazlama olasılığı 0,7 ve mutasyon olasılığı 0.001 dir.

Integer	Binary code	Integer	Binary code	Integer	Binary code
1	0001	6	0110	11	1011
2	0010	7	0111	12	1100
3	0011	8	1000	13	1101
4	0100	9	1001	14	1110
5	0101	10	1010	15	1111

Örnek:

- GA rasgele bir şekilde üretilen birler ve sıfırlarla altı adet 4-bit string ile doldurularak kromozomların başlangıç popülasyonunu oluşturur.
- Başlangıç popülasyonu Tablo 2 de gösterilen gibi görünebilir.

Tablo 2 Rasgele bir şekilde üretilen başlangıç kromozomlarının popülasyonu

Chromosome label	Chromosome string	Decoded integer	Chromosome fitness	Fitness ratio, %	Cumulative ratio
x_1	1100	12	36	16.5 %	16.5
x_2	0100	4	44	20.2 %	36.7
x_3	0001	1	14	6.4 %	43.1
x_4	1110	14	14	6.4 %	49.5
x_5	0111	7	56	25.7 %	75.2
x_6	1001	9	54	24.8 %	100.0

- Tablo 2 deki son kolon, popülasyonun toplam uygunluğuna bireysel kromozomların uygunluk oranını gösterir.
- Bu oran, eşleşme için kromozomların seçilme şansını tanımlar. Böylece, x_3 ve x_4 kromozomlarının çok düşük bir seçilme olasılığı varken, x_5 ve x_6 kromozomlarının yeterli şansı vardır.

Örnek: Seçim

- Eşleşme için bir kromozom seçiminde, rulet tekerleği seçimi tekniğini kullanarak, $[0,100]$ aralığında rasgele bir sayı üretilir ve kümülatif oranının rasgele sayıyı içerdiği kromozom seçilir.
- Bu, uygunluklarıyla orantılı olarak tekerlek üzerinde bir alana sahip her bir kromozomun yer aldığı bir tekerleğin dönmesi gibidir.
- Örneğimizde, biz altı kromozomlu bir başlangıç popülasyonuna sahibiz.
- Böylece, sonraki nesilde aynı popülasyon boyutunu kurmak için, altı rasgele sayı üretilecektir.(bu, rulet tekerleğinin altı kez dönmesi gibidir.)
- ilk iki dönme, x6 ve x2 ebeveyn olmak için seçebilir; ikinci dönme çifti, x1 ve x5 kromozomlarını seçebilir ve son iki dönüş x2 ve x5 kromozomlarını seçebilir.

Örnek: Çaprazlama

- Bir çift ebeveyn kromozom seçildikten sonra, çaprazlama operatörü uygulanır.
- Öncelikle çaprazlama operatörü rasgele bir şekilde, iki ebeveyn kromozomu kıracak ve bu noktadan sonra kromozomların parçalarını değiştirecek bir çaprazlama noktası seçer.
- Sonuç olarak iki yeni yavru oluşturulur.
- Mesela, iki kromozom x_6 ve x_2 ikinci genden sonra çaprazlanabilir. Her biri iki yavru üretir. Aşağıdaki gibi gösterilir:



- Ebeveynlerin tam kopyası olarak oluşturulur.
- 0.7 çaprazlama olasılığı genellikle iyi sonuç üretir.

Örnek: Mutasyon

- Doğada nadiren olan Mutasyon, gendeki bir değişikliği temsil eder.
- Bu, uygunlukta önemli bir gelişmeye neden olabilir. Fakat çoğunlukla, oldukça faydalı sonuçları vardır.
- Mutasyonun rolü, arama algoritmasının bir local optimuma takılmamasının garantisini sağlamaktır.
- Seçim sırası ve çaprazlama operatörleri herhangi bir homojen çözüm kümesinde durgunlaşabilir.
- Böyle şartlar altında, tüm kromozomlar özdeştir ve bu yüzden popülasyonun ortalama uygunluğu geliştirilemeyebilir.
- Çözüm sadece optimal (veya local olarak oldukça optimal) olmak görünebilir.
- Çünkü arama algoritması daha fazla ilerlemeyebilir. Mutasyon rasgele bir aramaya eşdeğerdir ve genetik farklılıkların kaybının korunmasında bize yardım eder.

Örnek: Mutasyon

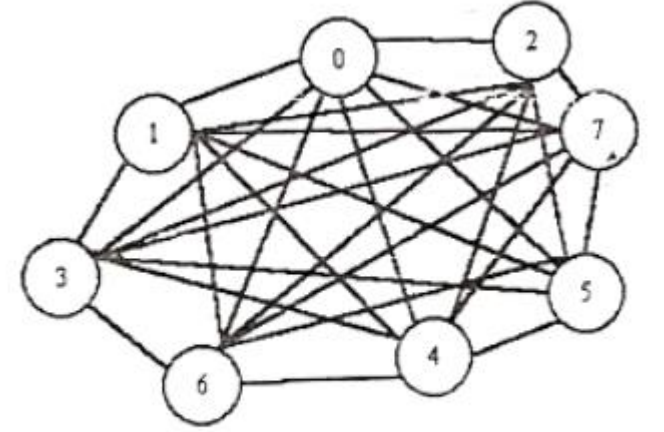
- Mutasyon operatörü bir kromozomda rasgele bir şekilde seçilen geni çevirir.
- Mesela, aşağıda gösterildiği gibi, x_1 'in ikinci geninde ve kromozom x_2 de üçüncü geninde mutasyona uğratılmış olabilir.

x_1	<table><tr><td>1</td><td>1</td><td>1</td><td>1</td></tr></table>	1	1	1	1	<table><tr><td>1</td><td>1</td><td>0</td><td>1</td></tr></table>	1	1	0	1
1	1	1	1							
1	1	0	1							
x_2	<table><tr><td>0</td><td>1</td><td>0</td><td>0</td></tr></table>	0	1	0	0	<table><tr><td>0</td><td>1</td><td>0</td><td>0</td></tr></table>	0	1	0	0
0	1	0	0							
0	1	0	0							
	Parents	Children								

- Mutasyon, bir kromozomda rasgele bir gende meydana gelebilir.
- Mutasyon olasılığı doğada oldukça küçüktür ve genetik algoritmalarda oldukça küçük tutulur.
- Tipik olarak 0.001 ile 0.01 aralığında.

Gezgin Satıcı Probleminin Genetik Algoritmalarla Uygulaması

- GA'ların uygulanmasına ilişkin incelenecek olan GSP'de 8 şehir olduğunu varsayalım.
- Hatırlatalım ki problemde amaç, satıcının her şehre yalnızca bir kere uğrayarak turunu en kısa yoldan tamamlamasıdır.
- Satıcının başladığı şehre geri dönme zorunluluğunun olmadığını varsayalım. (Yani burada amaç döngü değil, Hamilton yolunun bulunmasıdır.)
- Şehirler birbirine bağlı olmakta ve uzaklıklar verilmektedir bu şartlar içerisinde toplam alınan yolun minimum değerini bulmaya çalışacağız



Şekil 23: 8 Şehirli GSP Şeması

Şekil. 8 şehirli GSP şeması

Başlangıç popülasyonun oluşturulması:

- GSP“de 8 şehir olduğunu kabul etmiştik.
- Şehirleri “0” dan “7” ye kadar numaralandırdığımızı düşünürsek, her gen bir şehre karşılık gelecektir.
- Bu durumda kromozomlarımız ve başlangıç popülasyonumuz rastlantısal olarak aşağıdaki gibi oluşturulabilir.(Popülasyon büyüklüğü “5” olarak seçilmiştir.)
 - 10275463
 - 71302564
 - 24170536
 - 41653720
 - 60534217

Populasyon içindeki her kromozomun amaç fonksiyonunun değerinin hesaplanması:

- ilk kromozomumuza bakarsak, bu kromozomun amaç fonksiyonu değeri, şartta verilen şehirlerarası yol uzunluklarına bağlı olarak aşağıdaki şekilde hesaplanır:
- 10275463
 - 1"den 0"a 5 km
 - 2"den 7"ye 7 km
 - 5"den 4"e 3 km
 - 6"dan 3"e 5 km
 - 0"dan 2"ye 3 km
 - 7"den 5"e 4 km
 - 4"den 6"ya 6 km ise;
- Bu kromozomun gösterdiği yolun toplam uzunluğu 27 km, yani amaç fonksiyonun değeri "27" olacaktır.

Tekrar üretme, çaprazlama ve mutasyon operatörlerinin uygulanması:

- Goldberg probleminde uygulanan çaprazlama işlemini GSP'nin kromozomlarına uygularsak bazı güçlüklerle karşılaşabiliriz.
- Örneğin, rastgele bir şekilde üçüncü ve dördüncü kromozomları seçelim ve çaprazlama noktası 5 olsun.
- Bu durumda çaprazlama şu şekilde olacaktır.

Çaprazlama öncesi durum;

2 4 1 7 0 **5 3 6**

4 1 6 5 3 **7 2 0**

Çaprazlama sonrası durum;

2 4 1 7 0 **7 2 0**

4 1 6 5 3 **5 3 6**

- Görüldüğü gibi yeni oluşan kromozomlar, GSP problemi için olumlu bir sonuç vermez.
- Çünkü satıcımız, tüm şehirleri gezmeli ve her şehre yalnızca bir kez uğramalıydı, fakat çaprazlama sonrası kromozomlarda bazı şehirlere hiç gidilmemekte yada şehirlere uygun gelen düğümlere iki kere başvurulmaktadır.

Tekrar üretme, çaprazlama ve mutasyon operatörlerinin uygulanması

- Örneğin çaprazlama sonrası oluşan ilk kromozomumuza dikkat edecek olursak, 3., 5. ve 6. şehirlere hiç gidilmediğini ve 0., 2. ve 7. şehirlere iki defa uğrandığını gözlemleriz.
- Böyle bir olumsuz durumda ikinci bir işleme gerek duyulur.
- Yani kısıtlamaları sağlamayan uygunsuz kromozomların standartlaştırılması gerekmektedir.
- Standartlaştırma için aşağıdaki basit kural uygulanabilir: “Kromozom içinde
- tekrar eden ilk şehir ziyaret edilmeyen en küçük numaralı şehir ile değiştirilsin”.
- Bu standartlaştırma kuralı akla ilk gelen kurallardan biridir.

Tekrar üretme, çaprazlama ve mutasyon operatörlerinin uygulanması

- Çaprazlama ve standartlaştırma sonrası durum;
 - 34156720
 - 41027536
- Böyle bir durumda ikili düzende kodlanmış olan kromozomların daha avantajlı olabileceği görülmektedir.
- Ancak ikili düzende kodlama da kromozomların çok uzun olması, ya da problemin yapısına uygun olmaması gibi sorunlar çıkabilmektedir.
- Şimdi de Goldberg problemindeki gibi mutasyon operatörüne geçelim.
- Ancak bu rasgele seçilen kromozomlar üzerindeki iki genin yerlerini değiştirmekle gerçekleştirilir.
- Rastgele olarak seçtiğimiz 5. kromozomumuzun ikinci ve yedinci genlerini seçtiğimizi varsayalım.
- Bu durumda; mutasyondan önceki durum mutasyondan sonraki durum

60534217 67534210