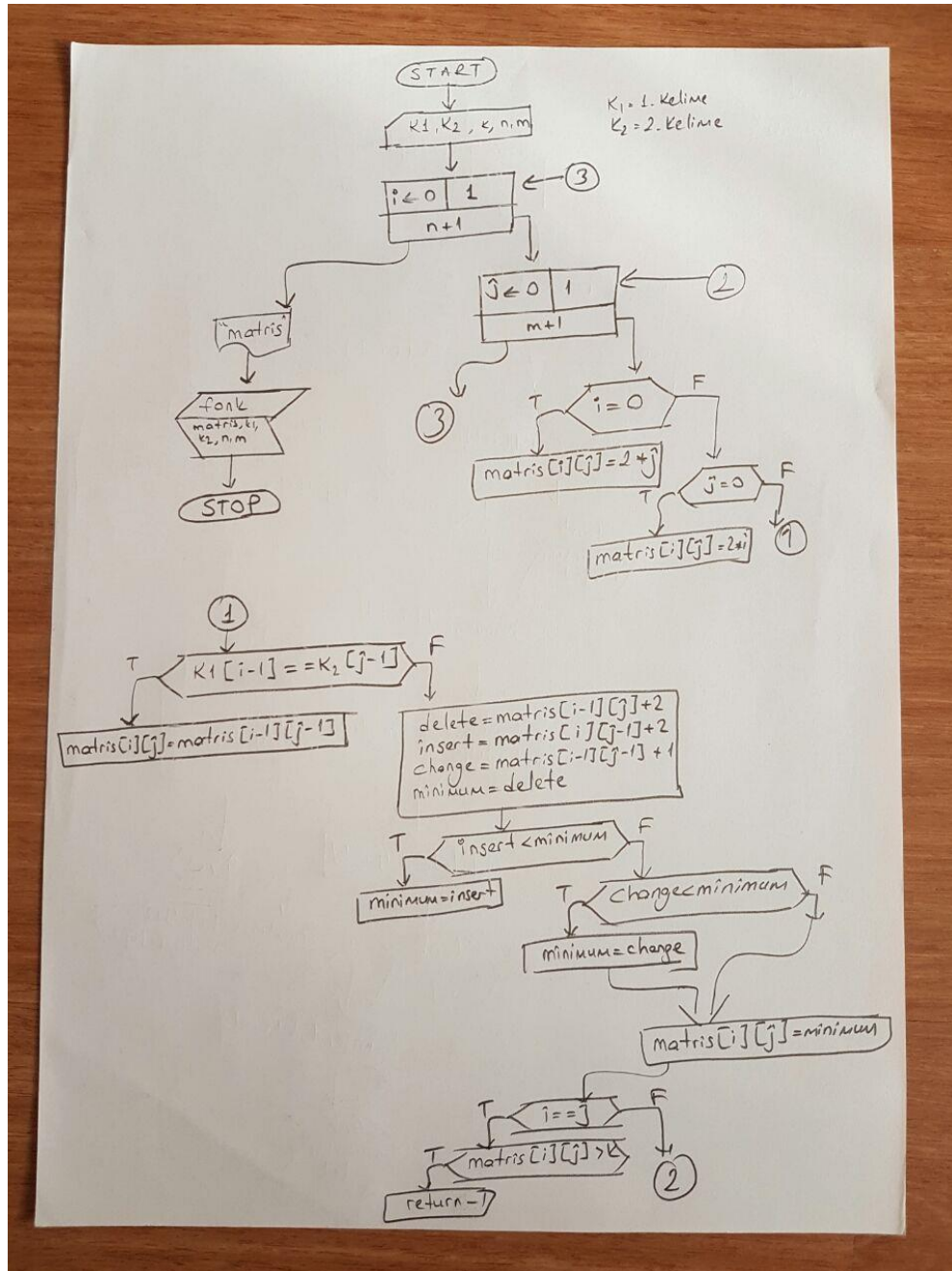


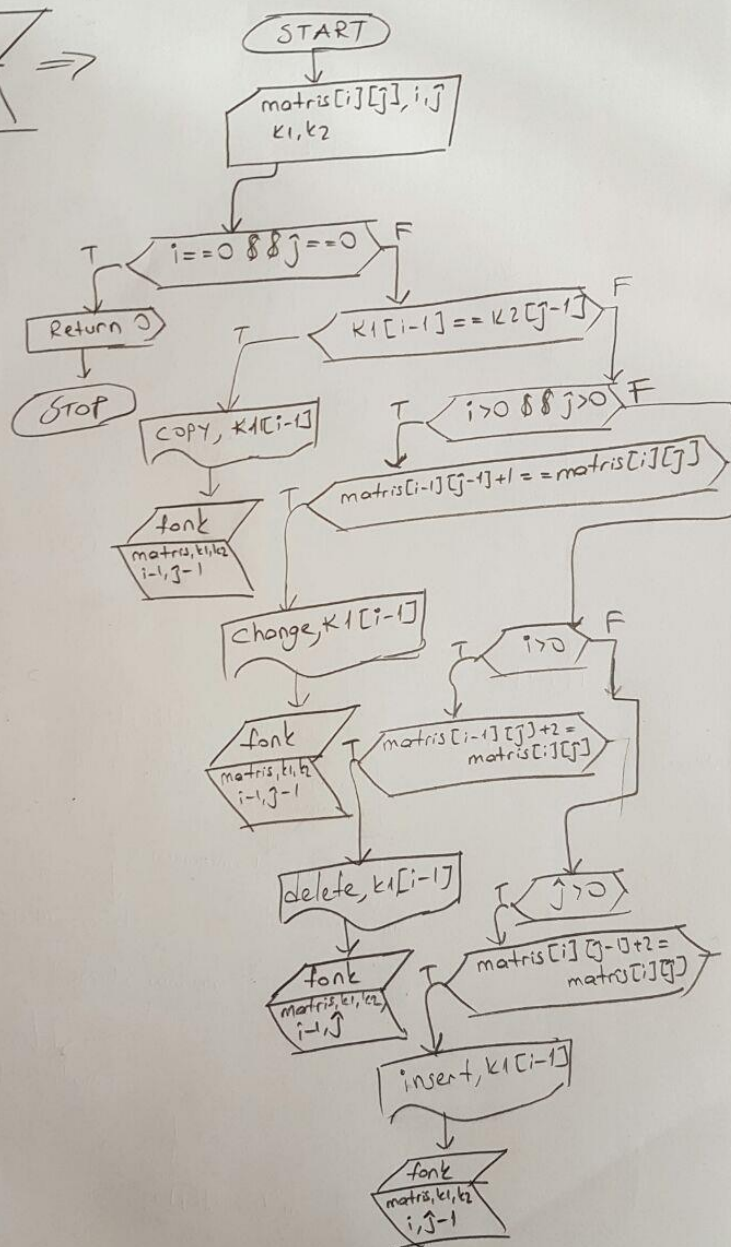
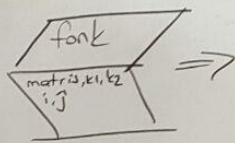
ALGORİTMA ANALİZİ 2 .ÖDEV

1).Yöntem

Bu ödevde iki genetik dizilimin en iyi hizalamasını Levenshtein Edit Distance yöntemi ile yapan bir algoritma tasarlanması istenmiştir. Bu yöntemde matris yapısı kullanılmaktadır. Matrisin(0..n) ve (m..0) gözleri ceza puanına göre ardışık olarak iklenir. Daha sonra verilen iki dizilim harf bazında kıyaslanarak belirlenen insert, delete, change işlemlerine göre ceza puanıyla birlikte matris doldurulmaya devam edilir. En sonda matrisin(n,m) lik gözü Edit Distance sayısını vermektedir. Yapılan işlemleri görebilmek için matrisin(n,m) lik gözünden itibaren Backtrace yöntemi uygulanır.

Algoritma Akış Diyagramı

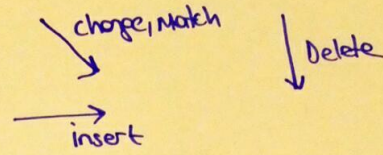




2). Uygulama

2- UYGULAMA

$$ED[i, j] = \min \begin{cases} ED[i-1, j-1] + 1 & \text{Change} \\ ED[i, j-1] + 1 & \text{insert} \\ ED[i-1, j] + 1 & \text{Delete} \\ ED[i-1, j-1] & \text{Match } x_i = y_j \end{cases}$$



a) ALGORITHM-ANALYSIS kelimeleri arasındaki mesafe 8'dir.

		A	N	A	L	Y	S	I	S
A	1	0	1	2	3	4	5	6	7
L	2	1	0	2	3	4	5	6	7
G	3	2	1	2	3	4	5	6	7
O	4	3	2	3	4	5	6	7	8
R	5	4	3	4	5	6	7	8	9
I	6	5	4	5	6	7	8	9	10
T	7	6	5	6	7	8	9	10	11
H	8	7	6	7	8	9	10	11	12
M	9	8	7	8	9	10	11	12	13

b) bazı depişim sekelleri:

• A-NALYSIS
ALGORITHM

match Delete Change Change Change Change Change

• ANALYSIS -
ALGORITHM

match change Delete

• A-NALYSIS -
A - - L G O R I T H M -

match insert insert copy change Delete copy change Delete

Analiz:	i	j	x _i	y _j	ED[i, j]	ED[i-1, j-1]	ED[i-1, j-1]+1	ED[i, j-1]+1	ED[i-1, j]+1
1	1	1	A = A		0	0			
1	2	2	A ≠ N		1		1	1	1
	3	3	A = A		2	ED[0, 2]	4	2+1	4
	4	4	A ≠ L		3		3+1	4+1	
	5	5	A ≠ Y		4		4+1	5+1	
	6	6	A ≠ S		5		5+1	6+1	
	7	7	A ≠ I		6				
	8	8	A ≠ S		7				

2	1	2	L ≠ A		1	1+1	2+1	0+1	
2	2	3	L ≠ N		1	0+1	1+1	1+1	
	3	4	L ≠ A		2	2+1	1+1	2+1	

Örnek1:

GGGAATCACGAGAGCAGACAGATCACACAGGTTTATGGGTTCTACGACGAGTGTTTA
GGGAATCATGAGAGCAGACGATCACACAAGTTTATGGTTTCTATGATGAATGTTTA

Edit Distance 8 dir.

1 farklı yolla hizalanabilir.

Örnek2:

AGGCTATCACCTGACCTCCAGGCCGATGCCC
TAGCTATCACGACCGCGGTTCGATTTGCCCGAC

Edit Distance 21 dir.

21 farklı yolla hizalanabilir.

3). Sonuç

K sayısı ile kısıtlama olmaksızın:

Worst Case: $m+n+m*n$

$m=n \rightarrow \text{Worst Case} = 2*n+n^2$

en yüksek dereceli terim alınır, katsayılar ihmal edilir.

O halde $\rightarrow n^2$

Worst Case----- $\rightarrow Q(n^2)$

K sayısı işin içine girerse.

Worst Case: $m+n+m*n+m*n$

$m=n \rightarrow \text{Worst Case} = 2*n+2*n^2$

en yüksek dereceli terim alınır, katsayılar ihmal edilir.

O halde $\rightarrow n^2$

Worst Case----- $\rightarrow Q(n^2)$

Bonus Soru

$\rightarrow A[n]$ ve $B[n]$ kelimeleri ve k değeri için değişim matrisinde ana diyagonal de $[0,0]$ 'dan en fazla kaç adım ilerlemesi gereklidir?

Matris doldurulurken ana diagonal üzerine gelinmesi kontrol edilir. Eğer ana diagonal üzerindeki herhangi bir sayı verilen 'K' değerinden büyük olduğu anda hesaplama durdurulmalıdır.