



KOCAELİ ÜNİVERSİTESİ
BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ
YAPAY ZEKA , 2021-2022

A* İLE ÇOKLU DİZİ HİZALAMASI PROJESİ

170202037

Arda Talu

170202064

Feridun Suay Bayar

180202087

Yağızhan Şimşek

200202133

Hami Çetinkaya

Özet–Projemizin amacı A* arama algoritmasını kullanarak bir çoklu dizi hizalama algoritması yapmaktır.

Anahtar Kelimeler –A*,Çoklu Dizi Hizalama,Gen

I. GİRİŞ

Programımız girilen dizileri a* ve needleman Wunsch algoritmalarını kullanarak sıralamaktadır

II. YÖNTEM

Programımızı C# programlama dili ile geliştirdik. Programımızda kısa yol algoritmaları temel olarak alınmıştır.

III. DENEYSEL SONUÇLAR

İlk olarak a* algoritmasının ne olduğunu araştırdık. A* algoritması bir gratta bir düğümden diğer düğüme giden en kısa yolu sezgisel bir şekilde hesaplar. Hesaplaması şu şekildedir: Düğümden gidilebilecek diğer düğümlere olan uzaklık ,+ diğer düğümlerin hedef düğüme olan kuş uçuşu. En kısa yolu bulur ve en kısa yolun olduğu düğüme hareket eder. Bu hareket hedef düğüme gidene kadar sürer.

Ardından dizi hizalamasının ne olduğunu araştırdık . Dizi hizalaması dizilerin indexlerindeki karakterleri

olabildiğince birbirine denk bir hale getirmeye çalışmaktadır. Peşine dizi hizalama algoritmasının nasıl birden fazla dizide uygulanabileceğini araştırdık. Burada karşımıza çıkan yöntem a* algoritmasını kullanarak bir merkez dizi seçmek ve sırası ile diğer dizileri merkez diziyeye göre sıralamak oldu . Bunun ardından ise sıraladığımız dizileri sırası ile ekrana girdik.

Program a* için kullanılacak olan yol ve uçuş matrislerini bir Excel dosyasından otomasyon şeklinde oluşturmaktadır.

	2	3	4	5	6	7	8	9	10	11	12		51	52	53	54	55	56	57	58	59	130
A	G	C	A	T	T	A	C	G	T			ASCATACGGT	51	0	-12	-18	-18	-21	-21	-18	-12	-7
A	C	C	A	A	G	T	T	A	C			ACCAAGTACC	52	-12	0	-12	-18	-15	-6	-12	-12	0
G	G	G	G	A	A	T	C	T	C			GGGGAACTCC	53	-18	-12	0	-21	-9	-18	-9	-21	-18
C	D	D	E	G	O	A	G	C				CDDEGGAGGC	54	-18	-18	-21	0	-18	-15	-21	-15	-12
G	A	A	G	G	T	C	A	G	A			GAAAGGTCAGA	55	-18	-15	-9	-18	0	-15	0	-15	-18
C	C	A	A	C	G	T	T	C	C			CCAAAGTCCG	56	-21	-9	-18	-15	-18	0	-6	-18	-12
G	C	A	G	G	T	C	T	C	G			GCAAGGTCCTG	57	-21	-12	-12	-21	-15	-6	0	-18	-18
A	A	T	C	G	G	C	A	C	A			AATCGGACAC	58	-18	-12	-21	-15	-18	-18	0	-6	-15
A	C	T	A	G	A	G	A	C				ACTAGAGAAC	59	-12	-15	-18	-18	-15	-18	-6	0	-12
C	E	T	G	T	G	A	T	G	A			CTETGATGAT	130	-7	-18	-18	-12	-18	-12	-18	-12	0

RE

SE

1

	2	3	4	5	6	7	8	9	10	11	12		51	52	53	54	55	56	57	58	59	130
2	C	C	A	A	T	T	A	C	G	T			51	0	-12	-18	-18	-21	-21	-18	-12	-7
3	C	C	A	A	T	T	A	C	G	T			52	-12	0	-12	-18	-15	-6	-12	-12	0
4	C	C	A	A	T	T	A	C	G	T			53	-18	-12	0	-21	-9	-18	-9	-21	-18
5	C	C	A	A	T	T	A	C	G	T			54	-18	-18	-21	0	-18	-15	-21	-15	-12
6	C	C	A	A	T	T	A	C	G	T			55	-18	-15	-9	-18	0	-15	0	-15	-18
7	C	C	A	A	T	T	A	C	G	T			56	-21	-9	-18	-15	-18	0	-6	-18	-12
8	C	C	A	A	T	T	A	C	G	T			57	-21	-12	-12	-21	-15	-6	0	-18	-18
9	C	C	A	A	T	T	A	C	G	T			58	-18	-12	-21	-15	-18	-18	0	-6	-15
10	C	C	A	A	T	T	A	C	G	T			59	-12	-15	-18	-18	-15	-18	-6	0	-12
11	C	C	A	A	T	T	A	C	G	T			130	-7	-18	-18	-12	-18	-12	-18	-12	0

Kullanıcının programı çalıştırdığında gördüğü çıktı girilen gen dizileri , seçilmiş olan merkez dizinin numarası ve gen dizilerinin sıralanmış hali olmaktadır.

```
CDDDEGGAGGC
GAAGGGTCAGA
CCAACGTTCCG
GCAGGGTCTCG
AATCGGCACAC
ACTAAGAGAAC
CGTGTGATGAT
Merkez dizimiz:2. dizidir
```

The aligned genes are :

```
A C C A A G T T A C _ _ C
A G C _ A _ T T A C G _ G T
G G G G A A _ T C T _ C C
_ C D D D E G _ G A G G C
G A _ _ A G G G T C A G A
_ C C A A C G T T _ C C G
_ G C A G G G T C T _ C G
A _ _ A T C G G C A C A C
A C T A A G A G A A C
_ _ C G T G T G A T G A _ T
```

IV. YALANCI KOD

-BAŞLA

-road,fly ve visit matrislerimi oluştur

-gen dizilerimi tanımla

-gen dizilerimi ekrana yazdır

- road,fly ve visit matrislerimi doldur

-road ve fly matrislerimi control et

{

-Eğer matris simetrik değil ise hata ver

-Simetrik ise devam et

}

-A* algoritmasını çalıştır

{

-Başlangıç düğümünden hedef düğüme gidebilmek için
sırası ile gidilebilecek yolları dene

-Optimum yolu bul

-Yolun maliyetini bul

}

-Her diziden diğer dizilere olan yolun maliyetlerine
bakarak en uygun diziyi merkez matrisi olarak belirle

-Merkez dizisinin numarasını ekrana yazdır

-Tüm dizileri merkez dizisine göre sırala

{

-Verilen iki dizinin indexlerindeki değerleri birbirine en
denk hale getir

-Dizileri ekrana yazdır

}
-BİTİR

V. SONUÇ

Bu projeyi geliştirirken A* algoritmasını ve Çoklu Dizi
Hizalamasını öğrendik ve bu alanda kendimizi geliştirdik.

VI. KAYNAKÇA

- [1] *Dizi hizalaması*. (2010, August 20). Vikipedi: Özgür Ansiklopedi. Retrieved April 26, 2022, from https://tr.wikipedia.org/wiki/Dizi_hizalamas%C4%B1
- [2] . *Çoklu Dizi hizalaması*. (n.d.). Download million images for free. https://stringfixer.com/tr/Multiple_sequence_alignment
- [3] *A Yıldız Arama Algoritması (A star search algorithm, a*)*. (2009, 2). Bilgisayar Kavramları. <https://bilgisayarkavramlari.com/2009/03/02/a-yildiz-arama-algoritmasi-a-star-search-algorithm-a/>
- [4] *A* arama algoritması*. (n.d.). Download million images for free. https://stringfixer.com/tr/A-star_algorithm
- [5] *Needleman-wunsch algoritması*. (2014, November 22). Vikipedi: Özgür Ansiklopedi. Retrieved April 26, 2022, from https://tr.wikipedia.org/wiki/Needleman-Wunsch_algoritmas%C4%B1