



S3. **Veri Tabanı Sistemi** (Database System) nedir? Kaç tür **Veri Tabanı Sistemi** vardır? Açıklayınız.

(10)

---

S4. Birbirleri ile ilişkili dört farklı türe ait diziler ATCC, ATGC, TTCG, and TCGG olarak verilmiştir. Bu türlerin aralarındaki farkların ölçüsü olarak Hamming uzaklığı kullanan basit bir kümeleme algoritması kullanarak bu türlerin ilişkisini veren phylogenetic ağacı oluşturunuz.

(10)

---

S5. Aşağıda verilen Perl “one-liner” ne yapar, açıklayınız.

(5)

```
$ perl -npe 'last if /\d{4}$/' embl.data
```

---

S6. Bu soruyu çözmek için [http://www.yildiz.edu.tr/~naydin/na\\_Intro2Bio.htm](http://www.yildiz.edu.tr/~naydin/na_Intro2Bio.htm) sayfasına gidip ilgili linkten gerekli dosyayı indirmeniz gerekmektedir.

(20)

---

**CEVAPLAR**

C1. Solution:  
 D=Asp=Aspartic acid = {GAU, GAC}  
 H=His=Histidine = {CAU, CAC}  
 K=Lys=Lysine = {AAA, AAG}

Possible RNA sequences	Possible DNA sequences
GAU CAU AAA	CTA GTA TTT
GAU CAU AAG	CTA GTA TTC
GAU CAC AAA	CTA GTG TTT
GAU CAC AAG	CTA GTG TTC
GAC CAU AAA	CTG GTA TTT
GAC CAU AAG	CTG GTA TTC
GAC CAC AAA	CTG GTG TTT
GAC CAC AAG	CTG GTG TTC

C2. A database system is a computer program (or group of programs) that provides a mechanism to define and manipulate one or more databases.

- Personal database systems: Designed to run on PCs
  - Access, Paradox, FileMaker, dBase
- Enterprise database systems: Designed to support efficient storage and retrieval of vast amount of data
  - Interbase, Ingres, SQL Server, Informix, DB2, Oracle
- Open source database systems: Free!!! (Linux!!!)
  - PostgreSQL, MySQL

(a)

	S	E	V	M	L
E		*			
V			*		
C					
L					*

(b)

	S	E	V	M	L
E		*			
V			*		
C					
L					*

Dizilim 1:  
 S E V M \_ L  
 \_ E V \_ C L

Dizilim 2:  
 S E V \_ M L  
 \_ E V C \_ L

(c)

	-	S	E	V	M	L
-	0	0	0	0	0	0
E	0	0	5	2	0	0
V	0	0	2	9	6	3
C	0	0	0	6	8	5
L	0	0	0	3	8	12

E V M L  
 E V C L  
 $5 + 4 + -1 + 4 = 12$

C4.

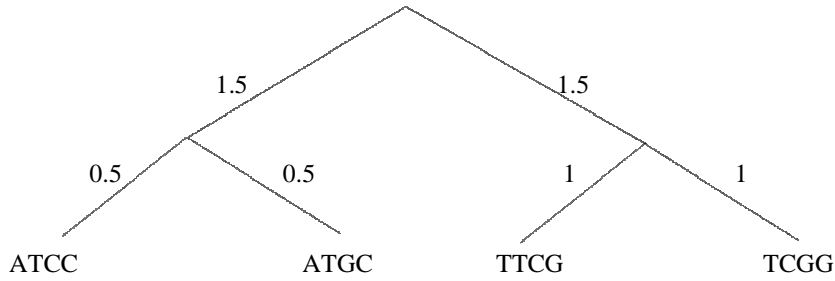
	ATCC	ATGC	TTCG	TCGG
ATCC	0	1	2	4
ATGC	1	0	3	3
TTCG	2	3	0	2
TCGG	4	3	2	0

	{ATCC,ATGC}	TTCG	TCGG
{ATCC,ATGC}	0	2.5	3.5
TTCG		0	2
TCGG			0

• İlk küme: { ATCC,ATGC}

Bir sonraki küme (cluster): {TTCG, TCGG}

Buna göre phylogenetic ağaç:



C5.

This is a Perl one-liner that prints only those lines from the embl.data disk-file that do not end in four digits.

C6. Çözümlerden biri aşağıdaki gibidir:

```
open (infile, "list.txt");
open (outfile, ">list_out.txt");
$cnt = 0;
while ($line = <infile>){
    if ( $line =~ m/Adaptive/ || $line =~ m/adaptive/) {
        $cnt++;
        $nlcon =~ s/ï>ï//g;
        $nlcon .= $line;
    }
}
print outfile "Total number of occurrences of the word "adaptive"= $cnt of occurrences.\n\n";
print outfile $nlcon;
```