

---

## BİL-211 Bilgisayar Programlama - I

### Ödev - 3

Veriliş Tarihi: 27.03.2018

Teslim Tarihi: 08.03.2018 (Saat : 23:59)

Teslim Şekli: Piazzadan duyurulacak.

**Kurallar:** Geç gönderilen ödevler kabul edilmez. Kopya kesinlikle yasaktır, kopya veren ve alan öğrenciler ödevden 0 alırlar ve ayrıca üniversite disiplin yönetmeliği kuralları bu öğrencilere uygulanır.

---

Bu ödevde metin dosyaları için basit bir sıkıştırma algoritması kodlayacaksınız. Sıkıştırma algoritmalarında amaç, dosyadaki karakterleri 8 veya 16 bitten daha az sayıda bitle gösterip sıkıştırılmış dosyaya kaydetmektir. Algoritma aşağıdaki adımlardan oluşuyor:

1. Dosyayı oku ve kaç farklı karakterden oluştuğunu belirle.
2. Her karakteri ikili düzende farklı bir sayıyla eşleştir. (Bunu yapabilmek için her karaktere kaç bitlik bir sayı gerek?)
3. Bütün dosyayı yukarıda belirlenen bitler şeklinde bir dizgiye (string) kaydet.
4. Bu dizginin bütün 8 karakterlik alt dizgilerini, onluk düzene çevirip bir byte dizisine at. (Yukarıda oluşan dizgi 8'in katları boyutunda olmayabilir!)
5. Bu byte dizisini (ve bazı ekta bilgileri) bir binary dosyaya yaz.

Algoritma sonunda oluşan dosya metin dosyasının sıkıştırılmış halidir. Örnek olarak, içeriği sadece aşağıdaki cümle olan metin dosyasına bakalım.

Programming is fun.

Dosyada 13 farklı karakter var: {P,r,o,g,a,m,i,n,s,f,u,, }. 13 farklı sayıyı ikili düzende gösterebilmek için en az 4 bite ihtiyaç vardır. Dolayısıyla karakter - sayı eşlemesi aşağıdaki gibi olacaktır:

P : 0000	r : 0001	o : 0010	g : 0011	a : 0100	m : 0101	i : 0110	n : 0111	s : 1000
f : 1001	u : 1010	. : 1011	:	1100				

Dolayısıyla dosyanın içeriği aşağıdaki dizgi ile gösterilebilir.

0000000100100011000101000101010101100111001111000110100011001001101001111011

Bu dizgiyi 8 karakterlik alt dizgiler halinde byte dizilerine atmanız gerekiyor. Dizgi toplam  $19 \times 4 = 76$  bitten oluşuyor ve 10 byte içinde saklanabilir (Son byte'a 4 bit kalıyor). Metin dosyasındaki bütün (farklı) karakterleri, onları kodlarını ve bu 10 byte'ı sırayla ikili dosyaya yazdığınız zaman sıkıştırma işlemi bitmiş olacak. Dosyanın bu ilk kısmına bu bilgileri nasıl yazacağımıza kendiniz karar vereceksiniz. Yapabildiğiniz kadar az yer kaplayacak şekilde yazmalısınız. (Bu örnek için sıkıştırılmış dosyanın metin dosyasından daha büyük olduğunu göreceksiniz, bu kadar küçük dosyalar için bu metot bir işe yaramaz, ama daha uzun metin dosyaları için algoritma sıkıştırma sağlayacaktır)

Tabi ki bu algoritma ile sıkıştırılmış bir dosyayı da programınızın açabilmesi gerekiyor.

Yazacağımız kod aşağıdaki gibi çalışmalıdır.

```
Komutu giriniz : Sıkıştır (S) | Aç (A) | Çık (Ç)
S
Sıkıştırılacak dosyayı giriniz : input.txt
input.txt.S dosyası oluşturuldu.
Komutu giriniz : Sıkıştır (S) | Aç (A) | Çık (Ç)
A
Açılacak dosyayı giriniz : input.txt.S
input.txt dosyası oluşturuldu.
Komutu giriniz : Sıkıştır (S) | Aç (A) | Çık (Ç)
Ç
```