

2015-2016 Bahar Yarıyılı
Veri Yapıları ve Algoritmalar 2. Ödevi

Konu : Huffman Kodlaması

Problem: Bu ödevde kayıpsız veri sıkıştırmak için kullanılan Huffman algoritması ile kodlanmış bir bilgiyi çözerek orjinal yazıyı elden bir algoritma tasarlanacak ve geliştirilecektir.

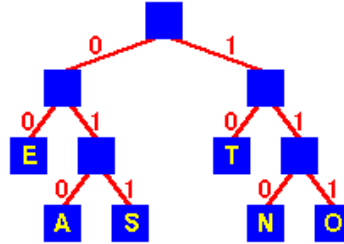
Ödev iki ana bölümden oluşmaktadır :

1. Huffman yöntemi ile **kodlanmış bilgiyi çözerek (decoding)** elde edilen yazıyı bir dosyaya yazınız.
2. Kodlama sonucunda **ortalama yer kazancını hesaplayınız.**

Kodlanmış bilgiyi çözmek için işlem adımları :

Kodlanmış bilginin hazırlanması:

1. Büyük harflerden oluşan bir yazı için **elle Huffman ağacını** ve **kodlanmış bilgiyi** hazırlayınız.
2. Huffman ağacına ait diziyi ve kodlanmış bilgiyi **2 ayrı dosyaya** yazınız.
3. Kodlanmış bilgiye ait bitler, **koda ait bütün bilgi aralarda boşluk olmadan ardarda olacak şekilde** saklanmalıdır. Eğer toplam bit sayısı 8'in katı değil ise 8'in katı olabilmesi için **sona yeteri kadar 0(sıfır) sayısı eklenmelidir.** Kod çözme işlemi sırasında, bu fazla sıfırların kod ile karışmasını engellemek için Huffman ağacını sakladığınız dosyaya toplam karakter sayısını da yazınız.
4. Ağacı gösteren diziyi hazırlarken, aşağıdaki örnekte olduğu gibi ağacın sol dallarının değerini 0, sağ dallarının değerini 1 yapınız. Ağacın **yapraklarında kodlanan karakterleri saklayınız.**



Kodlanmış bilginin çözülmesi:

1. Huffman ağacına ait diziyi ve kodlanmış bilgiyi **giriş bilgisi** olarak dosyalardan okuyunuz.
2. Kodlanmış bilgiyi unsigned char dizisi okuyunuz. Dizinin her elemanı kodun 8 bitlik bir parçasına aittir. Her 8 bitlik parçayı bit bit değerlendirmek için uygun işlem adımlarını yazınız.
3. Huffman ağacından faydalananarak kodu çözünüz.
4. Her adımda çözdüğünüz karakteri ve kod bilgisini ekrana yazdırınız.
5. Elde edilen yazıyı yeni bir dosyaya yazınız.

Kodlama sonucunda sıkıştırma oranının hesaplanması:

1. Yazıdaki her karakterin kullanım sayısını hesaplayınız.
2. Her karakterin kullanım sayısının o karakterin kodlandığı bit sayısı ile çarpımlarının toplamı sıkıştırılmış bilginin toplam bit sayısını **(n1)** verir. N1 değerini hesaplayınız.
3. Orjinal yazıda her karakter 8 bit ile gösterildiği için orjinal yazıdaki karakter sayısının 8 ile çarpılması orjinal yazının toplam bit sayısını **(n2)** verir. N2 değerini hesaplayınız.
4. Sıkıştırma oranı **CR = 1- (n1/n2)** değerini hesaplayınız.

Örnek: Yukarıda Huffman ağacı verilen yazıda bulunan karakterlerin kullanım sayısı A=16, E=22, N=10, O=18, S=14, T=20 olsun. Yazıdaki toplam karakter sayısı 100'dür. Sıkıştırılmış yazının toplam bit sayısı **n1**= 16*3 + 22*2 + 10*3+18*2+14*3+20*2=240 bit olur. Orjinal yazının bit sayısı **n2**= 100*8 = 800 bit'dir.

Sıkıştırma oranı **CR** = 1 – 240/800 = 0.7 olarak hesaplanır. Yazının sıkıştırma oranı %70'dir.

Teslim Edilecekler: Aşağıda verilen **bütün bilgileri içeren** tek bir doküman hazırlayınız. Raporun kapak sayfasına, dersin adı, öğrencinin ad, soyad ve numarası, ödev konusunu yazınız.

1. Yaptığınız çalışmayı **yöntem, uygulama ve sonuç** bölümlerinden oluşan bir raporda anlatınız.
 - **Yöntem Bölümü :** Problemin tanımını verip, gerçekleştirdiğiniz çözümü kısaca anlatıp, algoritmanıza ait akış diagramını çiziniz.
 - **Uygulama Bölümü :** Önerdiğiniz algoritmanın analizini yapınız. Analiz olarak sonucun ekran çıktısını vermeniz değil, derslerde yapıldığı gibi küçük bir örnek üzerinde ana değişkenlerin değişimini ve çözümün elde edilmesini adım adım göstermeniz gerekmektedir. **Bu ödev için, aşağıda verilen yazı parçasında:**
 - Huffman kodunu elle hesaplayıp rapora yazınız.
 - Ağacını elle oluşturarak rapora yazınız.
 - Kodlanmış yazıyı elle hazırlayarak rapora yazınız.
 - Kod çözümü işleminde ana değişkenlerin nasıl değiştiğini yazının ilk 3 karakteri için gösteriniz.
 - Sıkıştırma oranı hesabına ait bölümün analizini yapınız.

Analizde kullanılacak yazı parçası:

AAABBBBBBBCCDDDDDEEAACATTBBBBBFFCAAAABBBBBEEEEEEEECCCCCFFABDD
DDAAACCCCEEECCAADBBBBBAAAAACCCCAAAAAASSSSTTTTUUUUUUUUUU

- **Sonuç Bölümü :** Huffman kodlama yöntemi ile yazı sıkıştırma işleminin avantaj ve dezavantajlarını yorumlayınız.
2. Algoritmanızın C dilinde programını hazırlayarak dokümana ekleyiniz.

Teslim İşlemleri:

Ödevler **30 Mart 2016 Çarşamba günü** yapılacak laboratuvarda değerlendirilecektir. Laboratuvar günü ve saatleri Arş. Grv. Zeynep Banu Özger'in sayfasında ilan edilecektir. Ödev raporunuzu ve program kodunu **laboratuvar sunumunun yapılacağı günden 1 gün önce** Arş.Grv. Zeynep Banu Özger'e e-mail ile gönderiniz. Geç teslimler değerlendirilmeyecektir.