Predavanje I

Tags	Done

Izvorno kodiranje

Slova - skup simoba cini alfabet

Reci - konacni nizovi slova u alfabetu

Skup reci nad alfabetom je skup svih konacnih nizova simbola (a1..an)

Kodiranje se formalno definise kao funckija f koja slika odredjeni skup reci jednog alfabeta u drugi. Afabet izvora → Alfabet koda ($f : A' \to B'$)

Kod predstavlja skup svih slika. Elementi koda jesu kodne reci.

Blok fiksne duzine se odnosi kada se rec kodira sa odredjenim brojevem slova. Dok kod promenljive duzine se odnosi na zapis gde rec nije unapred predodredjena brojem slova koje treba da sadrzi.

Za predstavljanje koda se takodje moze koristiti kodno stablo.
Rec se dobija od korena stabla do odredjene grane.

Poruka se predstavlja nizovm kodnih reci.

```
x = v1,v2,…vn (v_k = f(r_k))
```

Kod koji omogucuje jednoznacno dekodiranje.
Klasa kodova koja omogucava jednoznacno kodiranje jeste
Prefiksno kodiranje.

Kod **V** je **prefiksni** ako ni jedna kodna rec iz **V** nije prefiks neke druge kodne reci iz **V**.

Prefiksni kod omogucava jednoznacno dekodiranje.

Dekodiranje → kod prefiksnih kodova dekodiranje se obavlja odmah po prijemu kodne reci. Zato se oni nazivaju trenutni kodovi.

Potrebno je da utvrdimo da li je nas kod jednoznacno dekodabilan. Algoritam koji nam govori da li je kod jednoznacno dekodabilan → Kriterijum jednoznacne

dekodabilnosti.
Kod V = {v_1…v_n} nad alfabetom B (V c B′) omogucuje
jednozacno dekodiranje akko je S podskup od V jednak praznim skupom, tj akko **ni jedan sufiks iz S nije kodna rec.**

Kada dobijemo kodnu reci koju dalje delimo i dobijanjem sufika S_1…S_n, ako naidjemo na kodnu rec koja pripada nekom od sufiksa to oznacava da kodnca rec nije jednoznacno dekodabilna.

Limitirajuci faktor za prefiksni kod je sto je to nerealno da ocekujemo da jedna kodna rec ima jedinstveni prefiks i da to nije deo nijedne druge kodne reci.

```
V = {01,010,110}
W = {1,101,010}
    S1  S2   S3    S4    S5    S6  S7
V 0   1,10 10
W 01   0   10  1,0 10,01 0,1 10,01

S = {0,1,10}
S presek V = prazan skup
-> kod je jednoznacno dekodabialn

S = {01,0,10,1}
S presek W nije prazan skup
-> kod nije jednoznacno dekodabilan
```

Kraftova nejednakost

Potrebno je da pokazemo da kodne reci odredjene duzine jednoznacno dekodabilne.
Cilj nam je dokazati da postoji prefisni kod sa istom duzinom reci.

Ukoliko imamo prefiksni kod **V** i imamo *b* reci u alfabetu koda, tada vazi da je *sum [b-n_i]* manja i jednaka od 1.

Ukoliko su dati brojeni n_1...n_a za koje vazi nejednakost, tada postoji prefiksni kod V tada postoji n_1...n_2 koje predstavljaju duzine odgovarajucih kodnih reci.

Analiticko svojstvo jedinstveno za prefiksne kodove.
Ako svaki prefiksni kod to zadovoljava, to znaci da je postoji prefiksni kod koji ima duzinu kodne reci.

```
V = {01,10,1110}

Pretpodstavljamo da su reci
sortirane po rastucoj duzini.

v1 = 01
v2 = 10
v3 = 1110

n je max duzina i formiramo skupove
tako da je svaka kodna rec duzine n

Bn1 = v1 = 01XX | 0100 | 0110 | 0101 | 0111
Bn2 = v2 = 10XX | 1000 | 1001 | 1010 | 1011
Bn3 = v3 = 1110

Primecujemo da nemamo dve iste reci.
Ukupan broj reci treba da bude manji
od vrednosti 2^n.

Obratna pretpostavka:
Pretpostavljamo da su brojevi n_1..n_a brojevi
koji zadovoljavaju kraftovu nejednakost i
du su poredjani u rastucem poretku
(bez gubitka opstosti => prvi korak je sortirati
n-ove ako nisu sortirani)

q_1 = 0, q_i = sum [b*n_a - n_k]

Potrebo je da formiramo kodne reci tako da
kod bude prefiksni i da svaka rec ima
odgovarajucu duzinu.
```

```
q_1 < q_2 < ... < q_n < 1

Imamo b simbola od kojih trebamo da fomriramo reci
b_1...b_na.
Formiramo sisteme sa osnovom b.
Cilj je j-ta cifra u sistemu sa osnovom b,
pri cemu cifre uzimamo do n_i
(n je restuci niz i zbog toga je to dovoljno)

qi = (C_1^i...C_ni^i) = sum C_j^ib^-j

Svaka rec ima n_i cifara, potrebno je
dokazati da je ovaj kod prefiksan.
Ne postoje dve kodne reci b_i,b_j
tako da su jedna drugoj prefiks.

q_j - q_i < b^-n_i

q_j - q_i = sum[b^n_k] >= b^-n_i
```