

Datastructuren

November 16, 2009

1. Implementeer Huffmans algoritme. Kies een geschikte manier om de bomen te representeren. Hoe bepalen we steeds de kleinste frequenties.
2. Een lange streng (string) bestaat uit vier characters A, C, G, T ; ze komen respectievelijk met een frequentie van 31%, 20%, 9% en 40% voor. Wat is de Huffman codering voor deze vier characters?
3. Veronderstel de symbolen a, b, c, d, e komen respectievelijk met frequentie $1/2, 1/4, 1/8, 1/16, 1/16$ voor.
 - a Wat is de Huffman codering van het alfabet?
 - b Als deze codering gebruikt wordt voor een bestand van 1.000.000 characters met de gegeven frequenties, wat is de lengte van het geëncodeerde bestand in bits?
 - c Als bovenstaande nu met frequenties: elk 20%
 - d Als bovenstaande nu met frequenties 5%, 15%, 20%, 25%, 35%
 - e Als bovenstaande nu met frequenties 40%, 20%, 20%, 10%, 10%
4. Geef de belangrijkste verschillen tussen binaire zoekbomen en bomen voor Huffman codes.
5. We gebruiken Huffmans algoritme om het alfabet $\{a, b, c\}$ met frequenties f_a, f_b, f_c te encoderen. Geef in elk van de volgende gevallen, of de een voorbeeld van frequenties f_a, f_b, f_c die aanleiding geven tot de gespecificeerde codering of leg uit waarom de codering niet verkregen kan worden (ongeacht de waarden voor de frequenties).
 - a Codering: $\{0, 10, 11\}$
 - b Codering: $\{0, 1, 00\}$
 - c Codering: $\{10, 01, 00\}$
6. We stellen de volgende codering voor: $a \rightarrow 00, b \rightarrow 001$ en $c \rightarrow 101$. Kan dit een (onderdeel van een) code zijn gevonden mbv. Huffmans algoritme? Laat zien dat een verzonden boodschap te ontcijferen is.
Lukt dit ook voor $a \rightarrow 00, b \rightarrow 001$ en $c \rightarrow 100$?
7. Geef een voorbeeld van een situatie waarin compressie van de bestanden zeer duidelijk wel mogelijk is maar waar Huffmans algoritme geen zoden aan de dijk zet.

8. Gegeven zijn 2^n letters met gelijke frequentie. Bewijs dat Huffmans algoritme een complete binaire boom oplevert.
9. Een *netwerk* is een gerichte graaf met gewichten en een speciale begin- en eindknoop. De takken van zo'n netwerk geven projecten aan en het bijbehorende gewicht de tijdsduur van het project. In zo'n netwerk komen geen kringen voor. De graaf geeft de afhankelijkheden tussen de projecten weer. Een project (i, j) kan pas gestart worden als alle projecten die eindigen in knoop i beëindigd zijn.
- Hoe bepalen we voor elke knoop i het (vroegste) moment waarop de projecten die uitgaan van knoop i gestart kunnen worden? We nemen hierbij aan dat de projecten die van de beginknoop uitgaan op tijdstip nul starten.
 - In sommige projecten kan vertraging ontstaan zonder dat dit effect heeft op de aanvangstijd van de andere projecten. Hoe kan bij elke knoop het laatste moment bepaald worden waarop alle inkomende projecten moeten eindigen zonder dat dit vertraging oplevert?
 - Geef aan hoe projecten gevonden kunnen worden waar beslist geen vertraging opgelopen mag worden zonder dat dit de totale benodigde tijd beïnvloedt. Deze projecten worden *kritiek* genoemd.
10. Voor een aantal projecten geven we *naam*, *tijdsduur* en projecten die *eerder uitgevoerd moeten zijn* vóórdat we het betreffende project mogen beginnen.
- | | | | | | | | | | | | | | | |
|-------|---|---|---|---|---|---|-----|---|---|---|---------|---|---|-------|
| Naam: | A | B | C | D | E | F | G | H | I | J | K | L | M | N |
| Duur: | 3 | 2 | 3 | 1 | 3 | 3 | 2 | 3 | 5 | 1 | 2 | 2 | 7 | 4 |
| Na: | - | A | B | C | B | B | F,H | A | A | H | D,E,G,J | I | I | K,L,M |
- Bepaal de totale tijdsduur benodigd voor het uitvoeren van de projecten. Welke projecten zijn 'kritiek' ?