

DATALOGISK INSTITUT, AARHUS UNIVERSITET

Science and Technology
EKSAMEN
Grundlæggende algoritmer og datastrukturer
Antal sider i opgavesættet (incl. forsiden): 11
Eksamensdag: Fredag d. 26. januar 2017, kl. 9.00 -11.00
Eksamenslokale: Åbogade 34 + Åbogade 36
Tilladte medbragte hjælpemidler: INGEN
Materiale der udleveres til eksaminanden: Opgavesæt

OPGAVETEKSTEN
BEGYNDER
PÅ NÆSTE SIDE

---o0o---

FADS 17 - Eksamen

Dette eksamenssæt består af to dele: (1) en mængde multiple-choice-opgaver som tæller 75%, og (2) en traditionel skriftlig opgave som tæller 25%. Multiple-choice opgaverne besvares på opgaveformuleringen **som afleveres**. Den traditionelle opgave besvares på separat papir som **afleveres sammen med opgaveformuleringen**.

For hver opgave er angivet opgavens andel af det samlede eksamenssæt.

For multiple-choice delen, har hvert delspørgsmål præcist et rigtigt svar. For hvert delspørgsmål, kan du vælge **max ét svar** ved at afkrydse den tilsvarende rubrik. Et delspørgsmål bedømmes som følgende:

- Hvis du sætter kryds ved det rigtige svar, får du 1 point.
- Hvis du ikke sætter nogen krydser, får du 0 point.
- Hvis du sætter kryds ved et forkert svar, får du $-\frac{1}{k-1}$ point, hvor k er antal svarmuligheder.

For en opgave med vægt $v\%$ og med n delspørgsmål, hvor du opnår samlet s point, beregnes din besvarelse af opgaven som:

$$\frac{s}{n} \cdot v\%$$

Bemærk at det er muligt at få negative point for en opgave.

For den traditionelle opgave bedes du forsøge at svare så kort og præcist som muligt, og undgå at skrive ting der ikke er relevante for opgaven. Du må svare på dansk eller engelsk. Husk at skrive dit navne herunder:

Navn og studienummer: _____

Multiple-choice del (75%)

Opgave 1 (6%) I denne opgave angiver $\log n$ 2-tals-logaritmen af n , \cdot er multiplikation og $/$ er division. Angiv om hvert af følgende udsagn er sande eller falske.

	Ja	Nej
$5n$ er $O(n^2)$?	<input checked="" type="radio"/>	<input type="radio"/>
n^3 er $O(5n^2)$?	<input type="radio"/>	<input checked="" type="radio"/>
$2n + 10$ er $O(n)$?	<input checked="" type="radio"/>	<input type="radio"/>
$n^{1/2}$ er $O(n^{1/3})$?	<input type="radio"/>	<input checked="" type="radio"/>
$n^{1/2} + n^2$ er $O(n)$?	<input type="radio"/>	<input checked="" type="radio"/>
$n^2/n^{1/2}$ er $O(n)$?	<input type="radio"/>	<input checked="" type="radio"/>
$n^{1/2} \cdot n^2$ er $O(n)$?	<input type="radio"/>	<input checked="" type="radio"/>
$1/n^2$ er $O(1/n)$?	<input checked="" type="radio"/>	<input type="radio"/>
$\log(4n)$ er $O(5 \log n)$?	<input checked="" type="radio"/>	<input type="radio"/>
$\log(n^4)$ er $O(\log n)$?	<input checked="" type="radio"/>	<input type="radio"/>
$(\log n)^2$ er $O(4 \log n)$?	<input type="radio"/>	<input checked="" type="radio"/>
$(\log n)^7$ er $O(n^{1/4})$?	<input checked="" type="radio"/>	<input type="radio"/>
$n^2/\log n$ er $O(n^{3/2})$?	<input type="radio"/>	<input checked="" type="radio"/>
2^n er $O(n^2)$?	<input type="radio"/>	<input checked="" type="radio"/>
4^n er $O(3^n)$?	<input type="radio"/>	<input checked="" type="radio"/>
3^n er $O(2^{2n})$?	<input checked="" type="radio"/>	<input type="radio"/>
$2^{\log n}$ er $O(n^2)$?	<input checked="" type="radio"/>	<input type="radio"/>
$2^{4 \log n}$ er $O(n^3)$?	<input type="radio"/>	<input checked="" type="radio"/>
n^2 er $O(3^{\log n})$?	<input type="radio"/>	<input checked="" type="radio"/>
n^3 er $O(8^{\log n})$?	<input checked="" type="radio"/>	<input type="radio"/>

Opgave 2 (6%) Betragt følgende max-heap, repræsenteret i et array A :

1	2	3	4	5	6	7	8	9	10
19	12	15	9	8	14	2	7	1	4

Angiv hvordan ovenstående max-heap ser ud efter indsættelse af elementet 13 (MAX-HEAP-INSERT(A , 13))

- | | | | | | | | | | | |
|----|----|----|---|---|----|---|---|---|----|----|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| 19 | 12 | 15 | 9 | 8 | 14 | 2 | 7 | 1 | 4 | 13 |
- | | | | | | | | | | | |
|----|----|----|---|----|----|---|---|---|----|----|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| 19 | 13 | 15 | 9 | 12 | 14 | 2 | 7 | 1 | 4 | 8 |
- | | | | | | | | | | | |
|----|----|----|----|---|----|---|---|---|----|----|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| 19 | 13 | 15 | 12 | 9 | 14 | 2 | 7 | 1 | 4 | 8 |
- | | | | | | | | | | | |
|----|----|----|---|---|----|----|---|---|----|----|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| 19 | 12 | 15 | 9 | 8 | 14 | 13 | 7 | 1 | 4 | 2 |

Opgave 3 (6%) Antag vi bruger MERGE-SORT til at sortere følgende array:

1	2	3	4	5	6	7	8
4	1	6	9	8	14	2	7

For hver af følgende par af lister, angiv om MERGE-SORT på ovenstående array vil kalde MERGE på sådant et par af lister som en del af sin udførsel.

Ja Nej

<table border="1"><tr><td>1</td><td>2</td><td>4</td><td>6</td></tr></table>	1	2	4	6	<table border="1"><tr><td>7</td><td>8</td><td>9</td><td>14</td></tr></table>	7	8	9	14	<input type="radio"/>	<input checked="" type="radio"/>
1	2	4	6								
7	8	9	14								

<table border="1"><tr><td>1</td><td>6</td></tr></table>	1	6	<table border="1"><tr><td>8</td><td>9</td></tr></table>	8	9	<input type="radio"/>	<input checked="" type="radio"/>
1	6						
8	9						

<table border="1"><tr><td>1</td><td>4</td><td>6</td></tr></table>	1	4	6	<table border="1"><tr><td>8</td><td>9</td><td>14</td></tr></table>	8	9	14	<input type="radio"/>	<input checked="" type="radio"/>
1	4	6							
8	9	14							

<table border="1"><tr><td>1</td><td>4</td><td>6</td><td>9</td></tr></table>	1	4	6	9	<table border="1"><tr><td>2</td><td>7</td><td>8</td><td>14</td></tr></table>	2	7	8	14	<input checked="" type="radio"/>	<input type="radio"/>
1	4	6	9								
2	7	8	14								

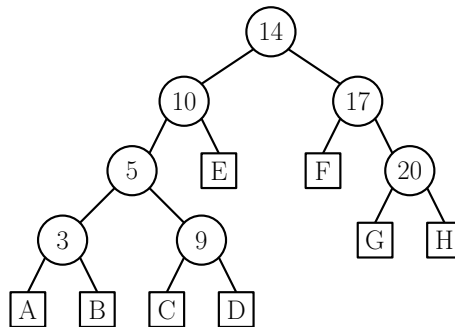
Opgave 4 (5%) Betragt RADIX-SORT anvendt på nedenstående liste af tal ($d = 5, k = 3$).

12220 21101 22100 01101 00002 22102

Angiv den delvist sortede liste efter at radix-sort har sorteret tallene efter de **tre** mindst betydende cifre.

- 00002 01101 12220 21101 22100 22102
- 00002 22100 01101 21101 22102 12220
- 22100 12220 21101 01101 00002 22102
- 00002 22100 21101 01101 22102 12220

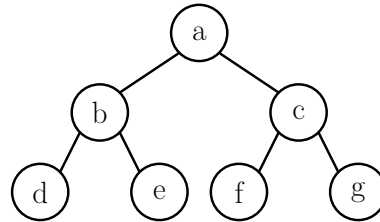
Opgave 5 (6%) Betragt følgende ubalancerede søgetræ:



Angiv i hvilke blade A-H elementerne 8, 19, 13, 15, 4 skal indsættes (det antages at før hver indsættelse indeholder træet kun ovenstående syv elementer).

	A	B	C	D	E	F	G	H
Insert(8)	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Insert(19)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
Insert(13)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Insert(15)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
Insert(4)	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

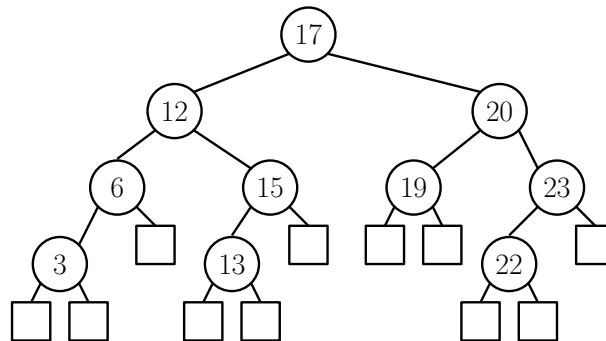
Opgave 6 (6%) Betragt følgende binære søgetræ indeholdende knuderne a, b, c, d, e, f, g :



Nøglerne hørende til knuderne vises ikke på figuren overfor, men det antages at det er et gyldigt søgetræ. Antag at vi nu laver en højre-rotation (RIGHT-ROTATE) på kanten imellem a og b . Angiv for hvert af følgende udsagn om det er sandt for søgetræet efter rotationen.

	Ja	Nej
e er venstre barn til a ($a.left = e$)	<input checked="" type="checkbox"/>	<input type="checkbox"/>
d er venstre barn til b ($b.left = d$)	<input checked="" type="checkbox"/>	<input type="checkbox"/>
c er højre barn til b ($b.right = c$)	<input type="checkbox"/>	<input checked="" type="checkbox"/>
b er venstre barn til c ($c.left = b$)	<input type="checkbox"/>	<input checked="" type="checkbox"/>

Opgave 7 (6%) For hver af nedenstående delmængder, angiv om følgende binære træ er et lovligt rød-sort træ hvis netop disse knuder farves røde



	Ja	Nej
3, 12, 13, 20, 22	<input checked="" type="checkbox"/>	<input type="checkbox"/>
3, 13, 22	<input checked="" type="checkbox"/>	<input type="checkbox"/>
3, 13, 12, 22	<input type="checkbox"/>	<input checked="" type="checkbox"/>
3, 6, 13, 15, 20, 22	<input type="checkbox"/>	<input checked="" type="checkbox"/>

Opgave 8 (5%) I følgende hashtabel er anvendt *linear probing* med hashfunktionen $h(k) = (k^2 + 2k) \pmod{11}$.

0	1	2	3	4	5	6	7	8	9	10
11		4	5					7		10

Angiv hvordan hashtabellen ser ud efter elementerne 3, 8 og 6 er blevet tilføjet (indsat i den rækkefølge, altså først 3, så 8 og tilsidst 6).

0	1	2	3	4	5	6	7	8	9	10
11		4	5	3	8	6		7		10

0	1	2	3	4	5	6	7	8	9	10
11	8	4	5	3				7	6	10

0	1	2	3	4	5	6	7	8	9	10
11	8	4	5	3	6			7		10

Opgave 9 (6%) Hver af følgende rekursionsligninger (recurrences) har basistilfældet $T(1) = 1$. Angiv for hver ligning, hvad løsningen er:

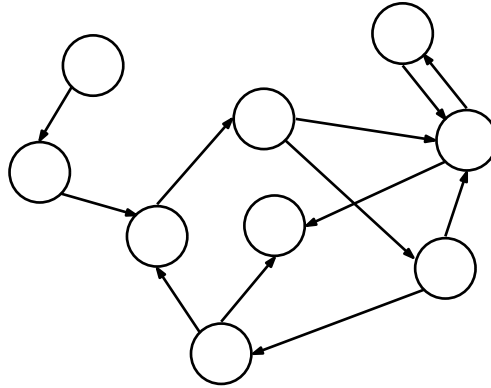
$T(n) = 4T(n/3) + n$ $\Theta(n)$ $\Theta(n \log n)$ $\Theta(n^{\log_3 4})$ $\Theta(n^{\log_4 3})$

$T(n) = 2T(n/2) + n$ $\Theta(n)$ $\Theta(n \log n)$

$T(n) = 3T(n/2) + n^2$ $\Theta(n^2)$ $\Theta(n^2 \log n)$ $\Theta(n^{\log_2 3})$ $\Theta(n^{\log_3 2})$

$T(n) = 8T(n/2) + n^3$ $\Theta(n^3)$ $\Theta(n^3 \log n)$ $\Theta(n^{\log_8 2})$

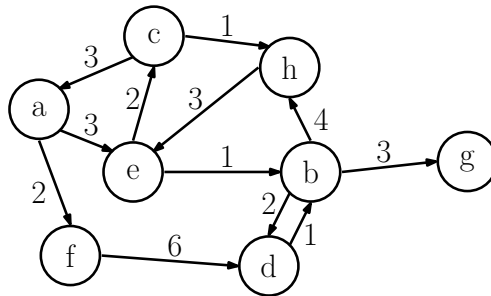
Opgave 10 (5%) Betragt følgende orienterede graf:



Angiv antallet af stærke sammenhængskomponenter (strongly connected components) i grafen

- 3
- 4
- 5
- 6

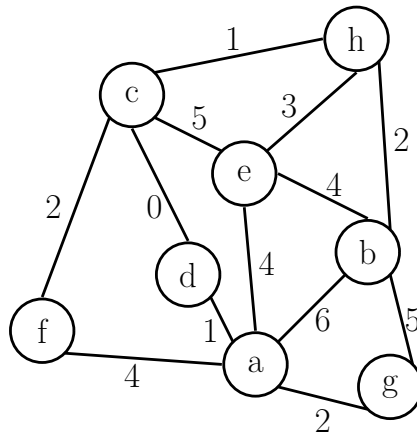
Opgave 11 (6%) Betragt følgende vægtede orienterede graf:



Antag vi kører Dijkstra's korteste veje algoritme (single source shortest paths SSSP), med knuden a som source. Efter 4 iterationer vil knuderne a , f , e og b være hevet ud af prioritetskøen (i den rækkefølge), og der er kaldt RELAX på deres udgående kanter. Angiv for hvert af følgende afstandsestimater, hvorvidt det svarer til det estimat Dijkstra's algoritme har på det tidspunkt, altså lige inden den 5. knude hives ud af prioritetskøen.

- | | Ja | Nej |
|-----------|----------------------------------|----------------------------------|
| $c.d = 5$ | <input checked="" type="radio"/> | <input type="radio"/> |
| $d.d = 8$ | <input type="radio"/> | <input checked="" type="radio"/> |
| $g.d = 7$ | <input checked="" type="radio"/> | <input type="radio"/> |
| $h.d = 6$ | <input type="radio"/> | <input checked="" type="radio"/> |

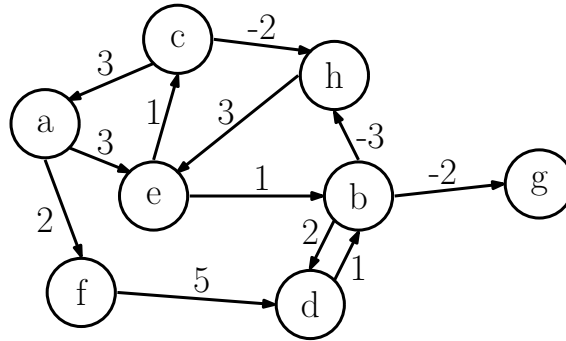
Opgave 12 (6%) Betragt følgende vægtede uorienterede graf:



Angiv for hver af følgende kanter, hvorvidt netop denne kant er med i det (unikke) minimum udspændende træ (minimum spanning tree MST) for grafen.

	Ja	Nej
{b,g}	<input type="radio"/>	<input checked="" type="radio"/>
{e,b}	<input type="radio"/>	<input checked="" type="radio"/>
{e,h}	<input checked="" type="radio"/>	<input type="radio"/>
{b,h}	<input checked="" type="radio"/>	<input type="radio"/>
{a,g}	<input checked="" type="radio"/>	<input type="radio"/>

Opgave 13 (6%) Betragt følgende vægtede orienterede graf:



Betragt den korteste vej i grafen ovenfor i blandt hver af følgende knudepar: (1) fra a til c , (2) fra h til d , (3) fra a til d , (4) fra b til a , (5) fra c til g . Angiv hvilken af de fem korteste veje, som er kortest:

- Fra a til c
- Fra h til d
- Fra a til d
- Fra b til a
- Fra c til g

Traditionel del (25%)

I ét-dimensionel k -means clustering, ønsker man at gruppere n tal i k clusters, således at summen af prisen for hvert cluster minimeres. Hvis et cluster C indeholder tallene x_1, \dots, x_ℓ , så er prisen, kaldet $\text{Cost}(C)$, for dette cluster defineret til:

$$\text{Cost}(C) := \sum_{i=1}^{\ell} \left(x_i - \frac{1}{\ell} \sum_{j=1}^{\ell} x_j \right)^2.$$

Vi ønsker altså at finde en opdeling af n tal i k mængder C_1, \dots, C_k , sådan at hvert tal er i præcist én mængde og hvor:

$$\sum_{i=1}^k \text{Cost}(C_i)$$

er mindst mulig i blandt alle sådanne opdelinger.

Hvis vi eksempelvis er givet tallene 1, 2, 6, 7, 21 for $k = 2$, så kunne en mulig opdeling være $C_1 = \{1, 2, 6\}$, $C_2 = \{7, 21\}$. Denne clustering/opdeling har prisen:

$$\begin{aligned} \sum_{i=1}^2 \text{Cost}(C_i) &= \\ \left(1 - \frac{1+2+6}{3}\right)^2 + \left(2 - \frac{1+2+6}{3}\right)^2 + \left(6 - \frac{1+2+6}{3}\right)^2 + \left(7 - \frac{7+21}{2}\right)^2 + \left(21 - \frac{7+21}{2}\right)^2 &= \\ (1-3)^2 + (2-3)^2 + (6-3)^2 + (7-14)^2 + (21-14)^2 &= \\ &= 112. \end{aligned}$$

En anden clustering ville være $C_1 = \{1, 2, 6, 7\}$ og $C_2 = \{21\}$. Denne clustering har prisen:

$$\begin{aligned} \sum_{i=1}^2 \text{Cost}(C_i) &= \\ \left(1 - \frac{1+2+6+7}{4}\right)^2 + (2-4)^2 + (6-4)^2 + (7-4)^2 + \left(21 - \frac{21}{1}\right)^2 &= \\ 3^2 + 2^2 + 2^2 + 3^2 + 0 &= \\ 26. & \end{aligned}$$

Den anden clusering er altså billigere.

Hvis vi lader x_1, \dots, x_n være n input tal, og hvis vi antager at de er sorteret således at $x_1 < x_2 < \dots < x_n$, så kan den optimale pris for en clustering af x_1, \dots, x_i ved at bruge k clusters beskrives som:

$$OPT(i, k) = \begin{cases} 0 & \text{if } k \geq i \\ \sum_{j=1}^i \left(x_j - \frac{1}{i} \sum_{\ell=1}^i x_\ell\right)^2 & \text{if } k < i \text{ and } k = 1 \\ \min_{j=1}^{i-1} \left(OPT(j, k-1) + \sum_{h=j+1}^i \left(x_h - \frac{1}{i-j} \sum_{\ell=j+1}^i x_\ell\right)^2 \right) & \text{if } 1 < k < i \end{cases}$$

For n input tal x_1, \dots, x_n , med $x_1 < x_2 < \dots < x_n$ ønsker vi altså at beregne værdien $OPT(n, k)$.

Opgave 14 (8%) Udfyld nedenstående tabel for $OPT(i, k)$ når input er tallene 1, 2, 6, 7 ($x_1 = 1, x_2 = 2, x_3 = 6, x_4 = 7$) og k går op til 2 (bemærk at indgangene ikke nødvendigvis er heltal):

$OPT(i, k)$	1	2
1	0	0
2	1/2	0
3	14	1/2
4	26	1

Opgave 15 (10%) Angiv en algoritme baseret på dynamisk programmering, der givet n tal x_1, \dots, x_n med $x_1 < x_2 < \dots < x_n$ og et positivt heltal k , beregner den mindst mulige pris for en k -means clustering af x_1, \dots, x_n i k clusters, dvs. beregner $OPT(n, k)$. Angiv algoritmens udførselstid.

Opgave 16 (7%) Udvid algoritmen til at rapportere de k clusters i en optimal clustering. Angiv algoritmens udførselstid.

Solution

Algorithm 1 KMEANS(x, K)

```
1:  $y$  = array of length  $n$ 
2:  $dp$  = table of size  $n \times K$ 
3:  $bt$  = table of size  $n \times K$ 
4: for  $i = 1$  to  $n$  do
5:   if  $i == 1$  then
6:      $y[i] = x[i]$ 
7:   else
8:      $y[i] = y[i - 1] + x[i]$ 
9:   for  $i = 1$  to  $n$  do
10:    for  $k = 1$  to  $K$  do
11:      if  $k \geq i$  then
12:         $dp[i][k] = 0$ 
13:      else if  $k == 1$  then
14:         $dp[i][k] = 0$ 
15:        for  $j = 1$  to  $i$  do
16:           $dp[i][k] = dp[i][k] + (x[j] - \frac{1}{i}y[i])^2$ 
17:        else
18:          for  $j = 1$  to  $i - 1$  do
19:             $v = dp[j][k - 1]$ 
20:            for  $h = j + 1$  to  $i$  do
21:               $v = v + (x[h] - \frac{1}{i-j}(y[i] - y[j]))^2$ 
22:            if  $j == 1$  or  $v < dp[i][k]$  then
23:               $dp[i][k] = v$ 
24:               $bt[i][k] = j$ 
25: return  $dp[n][K]$ 
```

Tid $O(n^3k)$ (men $O(n^4k)$ er også helt fint til eksamen)

Algorithm 2 BACKTRACK(x, K)

```
1:  $dp$  = table from KMEANS( $x, K$ )
2:  $bt$  = table from KMEANS( $x, K$ )
3:  $i = n$ 
4:  $k = K$ 
5: while  $k > 1$  do
6:    $j = bt[i][k]$ 
7:   print "cluster  $x_{j+1..i}$ "
8:    $i = j$ 
9:    $k = k - 1$ 
10: print "cluster  $x_{1..i}$ "
```

Tid $O(n + k)$