Question 1: Write your name and student number.

Solution: D. F. Search, 007

Question 2: Let $A[1 \ldots n]$ be an array storing $n$ numbers. We have seen algorithm \texttt{BuildHeap}(A) that rearranges the numbers in the input array $A$ such that the resulting array is a max-heap; see page 56 of my handwritten notes. This algorithm uses the \texttt{Heapify}-procedure as a subrouting; see page 53 of my handwritten notes. Consider the following variant of this algorithm:

\begin{algorithm}
\textbf{Algorithm} \texttt{BuildHeap}'(A):
\begin{algorithmic}
  \For{$i = 1$ to $\lfloor n/2 \rfloor$}
  \State \texttt{Heapify}(A, $i$)
  \EndFor
\end{algorithmic}
\end{algorithm}

Give an example of an array $A[1 \ldots n]$, where $n$ is a small integer (such as $n = 7$), which shows that algorithm \texttt{BuildHeap}' may not result in a max-heap.

Solution: We take the input array $A[1 \ldots 7] = [4, 6, 5, 3, 2, 7, 1]$. For this case, algorithm \texttt{BuildHeap}'(A) runs, in this order, \texttt{Heapify}(A, 1), \texttt{Heapify}(A, 2), and \texttt{Heapify}(A, 3).

The tree representation of the input array is the following:

\begin{tikzpicture}
  \node {4}
  \child {6}
  \child {3 \child {2} \child {7} \child {1}}
  \child {5}
\end{tikzpicture}

The call \texttt{Heapify}(A, 1) results in the following tree:

\begin{tikzpicture}
  \node {6}
  \child {4 \child {3} \child {2} \child {7} \child {1}}
  \child {5}
\end{tikzpicture}

The call \texttt{Heapify}(A, 2) does not change the tree. The call \texttt{Heapify}(A, 3) results in the following tree:

\begin{tikzpicture}
  \node {6}
  \child {4 \child {3} \child {2} \child {7} \child {1}}
  \child {5}
\end{tikzpicture}
This is not a max-heap, because element 7 is not at the root.

**Question 3:** Let $A[1 \ldots n]$ be an array storing $n$ pairwise distinct numbers, and let $k$ be an integer with $0 \leq k < n$. We say that this array is *k-sorted*, if for each $i$ with $1 \leq i \leq n$, the entry $A[i]$ is at most $k$ positions away from its position in the sorted order.

For example, a sorted array is 0-sorted. As another example, the array

$$A[1 \ldots 10] = [1, 4, 5, 2, 3, 7, 8, 6, 10, 9]$$

is 2-sorted, because each entry $A[i]$ is at most 2 positions away from its position in the sorted order. For $i = 3$, $A[3]$ is 2 positions away from its position, 5, in the sorted array. For $i = 9$, $A[9]$ is 1 position away from its position, 10, in the sorted array.

Describe an algorithm $\text{SORT}$ that has the following specification:

<table>
<thead>
<tr>
<th>Algorithm $\text{SORT}(A, k)$:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input:</strong> An array $A[1 \ldots n]$ of $n$ pairwise distinct numbers and an integer $k$ with $2 \leq k &lt; n$. This array is $k$-sorted.</td>
</tr>
<tr>
<td><strong>Output:</strong> An array $B[1 \ldots n]$ containing the same numbers as the input array. The array $B$ is sorted.</td>
</tr>
<tr>
<td><strong>Running time:</strong> Must be $O(n \log k)$.</td>
</tr>
</tbody>
</table>

Explain why your algorithm is correct and why the running time is $O(n \log k)$.

*Hint:* Use a min-heap of a certain size.

**Solution:** The approach is as follows:

- Let $H$ be the set consisting of the first $k + 1$ elements in the input array $A[1 \ldots n]$.
- Since the input array is $k$-sorted, the smallest element in the entire array $A[1 \ldots n]$ is the smallest element in the set $H$. We find the smallest element in $H$, delete it from $H$, and store it at $B[1]$.
- We add $A[k + 2]$ to the set $H$. Since the input array is $k$-sorted, the second smallest element in the entire array $A[1 \ldots n]$ is the second smallest element in the subarray $A[1 \ldots k + 2]$, which is the smallest element in the set $H$. We find the smallest element in $H$, delete it from $H$, and store it at $B[2]$. 


• We add $A[k + 3]$ to the set $H$. Since the input array is $k$-sorted, the third smallest element in the entire array $A[1 \ldots n]$ is the third smallest element in the subarray $A[1 \ldots k + 3]$, which is the smallest element in the set $H$. We find the smallest element in $H$, delete it from $H$, and store it at $B[3]$.

• We continue this process until $B[1 \ldots n - k - 1]$ stores, in sorted order, the $n - k - 1$ smallest element in the input array $A[1 \ldots n]$. At this moment, the set $H$ consists of the $k + 1$ largest elements in the input array $A[1 \ldots n]$. We add the elements of $H$ to the subarray $B[n - k \ldots n]$, one by one, from smallest to largest.

• How do we store the set $H$? We need the operations INSERT and EXTRACTMIN. This suggests that we store $H$ in a min-heap.

```
Algorithm Sort(A, k):
  Comment: Array $A[1 \ldots n]$ is $k$-sorted.
  Comment: The sorted numbers will be stored in array $B[1 \ldots n]$.
  initialize an array $H[1 \ldots k + 1]$;
  for $i = 1$ to $k + 1$
    do $H[i] = A[i]$
  endfor;
  BuildHeap($H$);
  for $i = 1$ to $n - k - 1$
    do $x = \text{EXTRACTMIN}(H)$;
      $B[i] = x$;
      $\text{INSERT}(H, A[k + 1 + i])$
  endfor;
  for $i = 1$ to $k + 1$
    do $x = \text{EXTRACTMIN}(H)$;
      $B[n - k - 1 + i] = x$
  endfor
```

Regarding the running time:

• Initializing the array $H$ takes $O(k)$ time, which is $O(n)$.

• The first for-loop takes $O(k)$ time, which is $O(n)$.

• The call to $\text{BUILDHEAP}(H)$ takes $O(k)$ time, which is $O(n)$.

• During the second for-loop, at any moment, the min-heap has size $k$ or $k + 1$, because we always delete the smallest element and then insert a new element. Each call to $\text{EXTRACTMIN}$ and $\text{INSERT}$ takes $O(\log k)$ time. The number of iterations of the second for-loop is $n - k - 1$, which is at most $n$. Thus, the total time for the second for-loop is $O(n \log k)$. 

3
• During the third for-loop, at any moment, the min-heap has size at most \( k + 1 \), because we only delete elements. Each call to \( \text{EXTRACTMIN} \) takes \( O(\log k) \) time. The number of iterations of the third for-loop is \( k + 1 \), which is at most \( n \). Thus, the total time for the third for-loop is \( O(n \log k) \).

• We conclude that the total running time is

\[
O(n) + O(n) + O(n) + O(n \log k) + O(n \log k) = O(n \log k).
\]

**Question 4:** You are given three beer barrels \( B_1, B_2, \) and \( B_3 \). Barrel \( B_1 \) has a capacity of 8 litres, barrel \( B_2 \) has a capacity of 5 litres, and barrel \( B_3 \) has a capacity of 3 litres.

At any moment, each barrel contains a given amount of beer (in litres). In one step, you can pour beer from one barrel, say \( B_i \), to another barrel, say \( B_j \). This step terminates at the moment when \( B_i \) becomes empty or \( B_j \) becomes full, whichever happens first.

To give some examples:

• If \( B_1 \) contains 6 litres of beer, \( B_2 \) contains 2 litres of beer, and \( B_3 \) contains 0 litres of beer, then we can pour the entire contents of barrel \( B_2 \) to barrel \( B_3 \). At the end of this step, \( B_1 \) contains 6 litres of beer, \( B_2 \) contains 0 litres of beer, and \( B_3 \) contains 2 litres of beer.

• If \( B_1 \) contains 3 litres of beer, \( B_2 \) contains 4 litres of beer, and \( B_3 \) contains 1 litre of beer, then we can pour 2 litres of beer from \( B_1 \) to \( B_3 \). At the end of this step, \( B_1 \) contains 1 litre of beer, \( B_2 \) contains 4 litres of beer, and \( B_3 \) contains 3 litres of beer.

**Decision problem:**

• Let \( b_1, b_2, \) and \( b_3 \) be integers such that \( b_1 \geq 0, b_2 \geq 0, 0 \leq b_3 \leq 3, \) and \( b_1 + b_2 + b_3 = 4 \). Similarly, let \( b'_1, b'_2, \) and \( b'_3 \) be integers such that \( b'_1 \geq 0, b'_2 \geq 0, 0 \leq b'_3 \leq 3, \) and \( b'_1 + b'_2 + b'_3 = 4 \).

• Initially, barrel \( B_1 \) is filled with \( b_1 \) litres of beer, barrel \( B_2 \) is filled with \( b_2 \) litres of beer, and barrel \( B_3 \) is filled with \( b_3 \) litres of beer.

• We want to decide whether or not it is possible to perform a sequence of steps that results in barrel \( B_1 \) having \( b'_1 \) litres of beer, barrel \( B_2 \) having \( b'_2 \) litres of beer, and barrel \( B_3 \) having \( b'_3 \) litres of beer?

(4.1) Formulate this as a problem on a directed graph. What are the vertices of the graph? What are the directed edges of the graph?

**Solution:** At any moment, the total amount of beer is equal to 4 litres. The current “state” is completely determined by the amount of beer in each barrel.

For any three integers \( b_1, b_2, \) and \( b_3 \) with \( b_1 \geq 0, b_2 \geq 0, 0 \leq b_3 \leq 3, \) and \( b_1 + b_2 + b_3 = 4, \) there will be one vertex, which we denote by \((b_1, b_2, b_3)\).
How many vertices are there? I am sure you remember from COMP 2804 that the number of integer solutions to the equation \( b_1 + b_2 + b_3 = 4 \) with \( b_1 \geq 0, b_2 \geq 0, \) and \( b_3 \geq 0, \) is equal to

\[
\binom{4 + 3 - 1}{3 - 1} = \binom{6}{2} = 15.
\]

Among these 15 solutions, there is one that cannot occur, namely \((0, 0, 4)\). Thus, our graph will have 14 vertices.

There is a directed edge from a source vertex to a target vertex if we can go in one step from the source vertex to the target vertex.

The decision problem becomes: given two vertices \( (b_1, b_2, b_3) \) and \( (b'_1, b'_2, b'_3) \), is there a directed path from the first vertex to the second vertex.
(4.2) Draw the entire graph.

Solution: As was to be expected, this is a pain to do. The two figures below show the graph. The first figure only shows the directed edges \((u, v)\), such that the reverse, i.e., \((v, u)\) is not an edge. The second figure shows all “symmetric” edges, i.e., those edges \((u, v)\) for which \((v, u)\) is also an edge. I made two figures, because everything in one figure is a complete mess.
(4.3) Assume that \((b_1, b_2, b_3) = (4, 0, 0)\) and \((b'_1, b'_2, b'_3) = (3, 1, 0)\). Use your graph to decide whether the answer to the decision problem is YES or NO.

**Solution:** We have to decide if there is a directed path from vertex \((4, 0, 0)\) to vertex \((3, 1, 0)\). The answer is YES. Here is one example of such a path:

\[
(4, 0, 0) \to (0, 4, 0) \to (0, 1, 3) \to (3, 1, 0)
\]

(4.4) Assume that \((b_1, b_2, b_3) = (4, 0, 0)\) and \((b'_1, b'_2, b'_3) = (2, 1, 1)\). Use your graph to decide whether the answer to the decision problem is YES or NO.

**Solution:** We have to decide if there is a directed path from vertex \((4, 0, 0)\) to vertex \((2, 1, 1)\). Since the vertex \((2, 1, 1)\) does not have any incoming edges, the answer is NO.

**Question 5:** Let \(G = (V, E)\) be an undirected graph. A *vertex coloring* of \(G\) is a function \(f : V \to \{1, 2, \ldots, k\}\) such that for every edge \(\{u, v\}\) in \(E\), \(f(u) \neq f(v)\). In words, each vertex \(u\) gets a “color” \(f(u)\), from a set of \(k\) “colors”, such that the two vertices of each edge have different colors.

Assume that the graph \(G\) has exactly one cycle with an odd number of vertices. (The graph may contain cycles with an even number of vertices.)

What is the smallest integer \(k\) such that a vertex coloring with \(k\) colors exists? As always, justify your answer.

**Solution:** Since the graph has an odd cycle, it cannot be colored using two colors. (This is the same as saying that the graph is not bipartite.)

We will show that the graph can be colored using three colors. Let \((v_1, v_2, \ldots, v_k, v_1)\) be the unique odd cycle in \(G\).

We remove one edge of this cycle, say \(\{v_1, v_2\}\) from \(G\), and denote the resulting graph by \(G'\). (We only remove this edge, we do not remove the vertices \(v_1\) and \(v_2\).)

We observe that the graph \(G'\) does not contain any odd cycle. Therefore, as was mentioned in lecture 10, \(G'\) is bipartite. Thus, we can split the vertex set \(V\) into two sets, say \(B\) and \(R\), such that every edge in \(G'\) has one vertex in \(B\) and the other vertex in \(R\).

We give every vertex of \(B\) the color blue, and every vertex of \(R\) the color red. This is a valid vertex coloring of \(G'\) using two colors. However, the vertices \(v_1\) and \(v_2\) have the same color. Thus, if we add the edge \(\{v_1, v_2\}\) to \(G'\), we do not get a vertex coloring of the original graph \(G\). We do get a vertex coloring of \(G\), by giving \(v_1\) a new color, say green.

**Question 6:** Let \(G = (V, E)\) be a directed graph, which is given to you in the adjacency list format. Thus, each vertex \(u\) has a list that stores all vertices of the set \(\{v : (u, v) \in E\}\).

The backwards graph \(G_b\) is obtained from \(G\) by replacing each edge \((u, v)\) in \(G\) by the edge \((v, u)\). In words, in \(G_b\), we follow the edges of \(G\) backwards.

7
Describe an algorithm that computes, in \( O(|V| + |E|) \) time, an adjacency list representation of \( G_b \). As always, justify your answer and the running time of your algorithm.

**Solution:** For each vertex \( u \), we write its adjacency list in \( G \) as \( A(u) \), and we write its adjacency list in \( G_b \) as \( A_b(u) \). The main observation is that for any two vertices \( u \) and \( v \),

\[ v \text{ is in } A(u) \text{ if and only if } u \text{ is in } A_b(v). \]

The algorithm does the following:

- For each vertex \( u \), initialize an empty list \( A_b(u) \).
- For each vertex \( u \), do the following:
  - For each vertex \( v \) in \( A(u) \), add the vertex \( u \) to the list \( A_b(v) \).

For the running time, initializing the lists \( A_b \) takes \( O(|V|) \) total time. The total time for the nested for-loops is proportional to

\[ \sum_{u \in V} (1 + |A(u)|) = |V| + |E|. \]

Thus, the total time for the entire algorithm is \( O(|V| + |E|) \).

**Question 7:** Let \( G = (V, E) \) be a directed acyclic graph, and let \( s \) and \( t \) be two vertices of \( V \).

Describe an algorithm that computes, in \( O(|V| + |E|) \) time, the number of directed paths from \( s \) to \( t \) in \( G \). As always, justify your answer and the running time of your algorithm.

**Solution:** We start by computing a topological sorting \( v_1, v_2, \ldots, v_n \) of the vertex set. Recall that for each edge \((v_i, v_j)\) in \( E \), \( i < j \). In other words, if we draw the vertices, in the given order, on a line, then all edges go from left to right.

If \( s \) is to the right of \( t \) in the topological sorting, then there is no directed path from \( s \) to \( t \). Thus, we assume that \( s \) is to the left of \( t \).

We may assume that \( s = v_1 \) and \( t = v_n \). (If, for example, \( s = v_7 \), then we can remove \( v_1, \ldots, v_6 \), and renumber the remaining vertices. Similarly, if, for example, \( t = v_{n-12} \), then we can remove \( v_{n-11}, \ldots, v_n \), and renumber the remaining vertices.)

We define \( P(1) = 0 \) and, for each \( i \) with \( 2 \leq i \leq n \), \( P(i) \) to be the number of directed paths from \( s \) to \( v_i \) in \( G \). Our task is to compute \( P(n) \).

For each \( i \), let \( IN(i) \) be the set of indices \( j \) such that \((v_j, v_i)\) is an edge in \( E \). Note that \( j < i \) for each such edge. The main observation is that

\[ P(1) = 0 \]
and for each \( i \) with \( 2 \leq i \leq n \),
\[
P(i) = \sum_{j \in \text{IN}(i)} P(j).
\]
This suggests that we can compute \( P(n) \) (this is the number we have to compute), by computing, in this order, \( P(0), P(1), P(2), \ldots, P(n) \).

The algorithm does the following:

- Compute a topological sorting \( v_1, v_2, \ldots, v_n \) of the vertex set \( V \). We have seen in class that this can be done in \( O(|V| + |E|) \) time.
- Use Question 6 to compute the list of incoming edges \( \text{IN}(i) \) for each vertex \( v_i \). This takes \( O(|V| + |E|) \) time.
- Initialize \( P(1) = 0 \). This takes \( O(1) \) time.
- For \( i = 2, 3, \ldots, n \), do the following:
  - Initialize \( P(i) = 0 \);
  - For each index \( j \) in \( \text{IN}(i) \), set
    \[
P(i) = P(i) + P(j).
    \]
  - This takes time
    \[
    O \left( 1 + \sum_{i=2}^{n} (1 + |\text{IN}(i)|) \right),
    \]
    which is \( O(|V| + |E|) \).
- Return \( P(n) \). This takes \( O(1) \) time.

The total running time of the algorithm is \( O(|V| + |E|) \).