Assignment One

Josh Seligman

joshua.seligman1@marist.edu

September 8, 2022

1 Singly Linked List

1.1 The Data Structure

A singly linked list is comprised of nodes which contain some form of data as well as a pointer to the next element within the list. As shown in Figure 1.1, the final node has a next of **null**, which marks the end of the list. In order to access a particular element, one has to start at the beginning and traverse through the list until the desired node is found. This causes data access to be on the magnitude of $O(n)$ as the time required to find an element has a linear relationship with the size of the list.

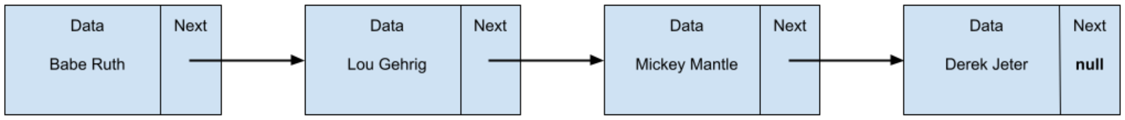

Figure 1.1: Example singly linked list of 4 Yankees legends.

1.2 Benefits of a Singly Linked List

1.2.1 Size

As previously mentioned, the last node within a linked list has a next of **null**. This characteristic enables linked lists to have no size restrictions barring memory capacity. As a result, this feature makes linked lists preferred over arrays, which have a fixed length, when the size of the data is frequently changing and has an unkown maximum. For instance, as demonstrated in Section 5.4 starting on line 10 within *main.cpp*, the size of the linked list is only limited by our needs and, if needed, more nodes are able to easily be added to the list with their creation as done on lines 13-15 and linking as shown on lines 18 and 19. On the other hand, if the list was made with an array, the size of the array would have to be provided at the time of the creation of the array, and it would not be easy to change the size if additional data have to be added to the array.

1.2.2 Data Type Flexibility

Linked lists do not have to be restricted to be able to store a specific data type. Instead, with the use of generics (C++ templates), the definition of a node is independent of the data type that the user wants to store within the linked list. This provides flexibilty and reusability for many use cases. As demonstrated in Section 5.1 in *node.h*, the definition of a node uses a generic T as the type of data being stored, which prevents any assumptions of the data and ensures compatibility with all data types. However, due to how the C++ linker works and to prevent all the code from being written within a single header file, the allowed types have to be stated on lines 13 and 14 of *node.cpp*. This is a C++ specific issue and is not present in other languages such as Java. Regardless, although they have to be specified for C++, any data type can still be stored within a node and a linked list. A demonstration of the user defining which data type is stored in a node is in Section 5.4 on lines 13-15 within *main.cpp*. Instead of the Node class defining the data type, the user is able to specify the type of data they want to store, which is a string in this situation but can be anything they want.

2 Problem One

2.1 The Data Structure

An n-element array of integer pairs: ($currentCount$, $predecessorSum$) will, once initialized in the preprocessing step, support Member, Less, and Range operations as described in Section 2.3 in $O(1)$ (constant) time.

For illustration, consider the case where $S = (1, 2, 3, 4, 7, 7, 7, 7, 7, 8, 9, 10, 10, 10, 10, 10, 10, 14, 16)$.
Here we have $m = 19$ elements in S ranging in value from 1 through $n = 16$. The array for S after preprocessing is given in Figure 2.1.

Figure 2.1: Example array built from 19 values, 10 of them unique.

2.2 Preprocessing

There are two preprocessing steps, Tally and Predecessor Sum, each of which is $O(m+n)$ as we will see in Equations 2.1 and 2.2 below. This makes the overall performance $O(m+n)$ because $2 \cdot O(m+n) =$[1] $O(m+n)$.

2.2.1 Pass one: Tally

For each element i in S we'll store its number of occurrences at $A[i].currentCount$.

```
Data: collection $S$
Result: array $A$ containing tallys for each element $i$ in $S$
allocate $A[n]$;
for $j \leftarrow 1$ to $n$ do
    $A[j].currentCount \leftarrow 0$;
end
foreach element $i$ in $S$ do // There are $m$ elements in $s$.
    $A[i].currentCount \leftarrow A[i].currentCount + 1$;
end
```

Algorithm 1: Tally

[1] This is an abuse of the notation, but if it's okay in the CLRS [?] book I hope it's okay here.

Asymptotic Analysis
We are given m integer values (the size of the collection), each in the range $[1..n]$ (the size of the domain) and want to iterate over them, computing the tally for each. Consider Algorithm 1. Line 1 executes in constant time. Lines 2 through 4 execute in $O(n)$ time because we are iterating from 1 to n. Lines 5 through 7 execute in $O(m)$ time because we are iterating over all the elements of S, of which there are m.

So, for pass one, we have :

$$\begin{aligned} O(pass1) &= constant + O(n) + O(m) \\ &= O(n) + O(m) \\ &= O(m+n) \end{aligned} \tag{2.1}$$

2.2.2 Pass two: Predecessor Sum

Once the tally is done we need to make another pass over A to compute, for each $A[i] : i > 1$, the sum of all of its predecessors $(A[1]..A[i-1])$ and store that in $A[i].predecessorSum$.

```
  Data: array $A$ containing tallys for each element $i$ in $S$
  Result: a new and improved array $A$, now with the predecessor sums
1 $A[1].predecessorSum \leftarrow 0$;
2 if $n > 1$ then
3     for $k \leftarrow 2$ to $n$ do
4         $A[k].predecessorSum \leftarrow A[k-1].predecessorSum + A[k-1].currentCount$;
5     end
6 end
```

Algorithm 2: Predecessor Sum

Asymptotic Analysis
Looking at Algorithm 2, line 1 is assignment, a constant time operation. If lines 3 through 5 execute at all, they do so in $O(n)$ time because we make $n-1$ iterations over line 4, which is assignment and array lookups, both constant time operations. For pass two, we have :

$$\begin{aligned} O(pass2) &= constant + O(n) \\ &= O(n) \\ &= O(m+n) \text{ which is ok so long as } m > 0 \end{aligned} \tag{2.2}$$

2.3 Operations

2.3.1 Member

```
  Input: parameter $i$
  Data: a new and improved array $A$, replete with tallys and predecessor sums
  Output: $True$ if $i$ exists in $S$, $False$ otherwise
1 if $(i \geq 1) \wedge (i \leq n)$ then
2     return $(A[i].currentCount > 0)$;
3 else
4     return $False$;
5 end
```

Algorithm 3: Member

Asymptotic Analysis
Looking at Algorithm 3, line 1 consists of comparisons, which are constant time operations. Line 2 is an array lookup and a comparison, both constant time operations. The remaining parts of the algorithm (including

the rest of line 2) are for program control, and not considered in this analysis. Since all parts of the algorithm execute in constant time, the whole thing executes in constant time and is therefore $O(1)$.

2.3.2 Less

```
  Input: parameter $i$
  Data: a new and improved array $A$, replete with tallys and predecessor sums
  Output: The number of elements in $S$ that are strictly less than $i$.
1 if $(i \geq 1) \wedge (i \leq n)$ then
2     return $A[i].predecessorSum$;
3 else
4     return 0;
5 end
```

Algorithm 4: Less

Asymptotic Analysis
Looking at Algorithm 4, line 1 consists of comparisons, which are constant time operations. Line 2 is an array lookup, a constant time operation. The remaining parts of the algorithm (including the rest of line 2) are for program control, and not considered in this analysis. Since all parts of the algorithm execute in constant time, the whole thing executes in constant time, and is therefore $O(1)$.

2.3.3 Range

```
  Input: parameters $i$ and $j : i \leq j$
  Data: a new and improved array $A$, replete with tallys and predecessor sums
  Output: The number of elements in $S$ that are in the range $[i..j]$.
1 if $(i \geq 1) \wedge (i \leq n) \wedge (j \geq 1) \wedge (j \leq n) \wedge (i \leq j)$ then
2     return $(A[j].currentCount + A[j].predecessorSum - A[i].predecessorSum)$;
3 else
4     return 0;
5 end
```

Algorithm 5: Range

Asymptotic Analysis
Looking at Algorithm 5, line 1 consists of comparisons, which are constant time operations. Line 2 consists of array lookups, addition, and subtraction, all constant time operations. The remaining parts of the algorithm (including the rest of line 2) are for program control, and not considered in this analysis. Since all parts of the algorithm execute in constant time, the whole thing executes in constant time, and is therefore $O(1)$.

3 Problem Two

4 Problem Three

5 Appendix

5.1 Singly Linked List

node.cpp

```cpp
#include <string>

#include "node.h"

template <typename T>
Node<T>::Node(T initialData) {
    // Initialize the node with the data and without a next node in the linked list
    Node::data = initialData;
    Node::next = nullptr;
}

// Define acceptable data types that the Node can accept for the template
template class Node<std::string>;
template class Node<char>;
```

node.h

```cpp
#pragma once

// Node represents an item within a singly linked list and can store data of a given type
template <typename T>
class Node {
    public:
        // A node has the data it is storing (of a type defined by the user)
        // and a pointer to the next node
        T data;

        // The pointer uses the template to make sure all elements of the linked list
        // store the same data type
        Node<T>* next;

        // Nodes will be instantiated with some data and not have a next node
        Node(T initialData);
};

// Super helpful resource on templates for c++
// https://isocpp.org/wiki/faq/templates#separate-template-fn-defn-from-decl
```

5.2 Stack

stack.cpp

```cpp
#include <string>

#include "stack.h"
#include "node.h"

// Instantiate the stack with the top pointing to nothing
template <typename T>
Stack<T>::Stack() {
    top = nullptr;
}

template <typename T>
Stack<T>::~Stack() {
    // Since the nodes were created on the heap, we have to
    // make sure everything is cleared from memory
    while (!isEmpty()) {
        pop();
    }
}

// Creates a new node and adds it to the stack
template <typename T>
void Stack<T>::push(T newData) {
    Node<T>* newNode = new Node(newData);
    // Set the next first so we do not lose the rest of the stack
    newNode->next = top;
    top = newNode;
}

// Removes the top node from the stack
template <typename T>
T Stack<T>::pop() {
    if (isEmpty()) {
        // Throw an exception if the stack is already empty
        throw std::invalid_argument("Tried to pop from an empty stack.");
    } else {
        // We need to collect the data in the node before removing it from the stack
        Node<T>* topNode = top;
        T topData = topNode->data;
        top = top->next;

        // Since the node was created on the heap, we have to free it from memory
        delete topNode;
        return topData;
    }
}

// Checks to see if the stack is empty or not
template <typename T>
bool Stack<T>::isEmpty() {
    return top == nullptr;
}

// Define acceptable data types that the Stack can accept for the template
template class Stack<std::string>;
template class Stack<char>;
```

stack.h

```cpp
#pragma once

#include "node.h"

template <typename T>
class Stack {
private:
    // Top points to the top of the stack
    Node<T>* top;
public:
    // We need a constructor and destructor
    Stack();
    ~Stack();

    // Push adds a new element to the stack
    void push(T newData);

    // Pop removes the top element from the stack
    T pop();

    // isEmpty checks to see if the stack is empty
    bool isEmpty();
};
```

5.3 Queue

queue.cpp

```
#include <string>

#include "queue.h"
#include "node.h"

// Instantiate the queue with the head pointing to nothing
template <typename T>
Queue<T>::Queue() {
    head = nullptr;
}

template <typename T>
Queue<T>::~Queue() {
    // Since the nodes were created on the heap, we have to
    // make sure everything is cleared from memory
    while (!isEmpty()) {
        dequeue();
    }
}

// Creates a new node and adds it to the queue
template <typename T>
void Queue<T>::enqueue(T newData) {
    Node<T>* newNode = new Node(newData);

    if (isEmpty()) {
        // Immediately set the head to be the new node if we are empty
        head = newNode;
    } else {
        // Traverse to the back of the queue
        Node<T>* cur = head;
        while (cur->next != nullptr) {
            cur = cur->next;
        }
        // Insert the new node in the back of the queue
        cur->next = newNode;
    }
}

// Removes the front node from the queue
template <typename T>
T Queue<T>::dequeue() {
    if (isEmpty()) {
        // Throw an exception if the queue is already empty
        throw std::invalid_argument("Tried to dequeue from an empty queue.");
    } else {
        // We need to collect the data in the node before removing it from the queue
        Node<T>* frontNode = head;
        T frontData = frontNode->data;
        head = head->next;

        // Since the node was created on the heap, we have to free it from memory
        delete frontNode;
        return frontData;
    }
}

// Checks to see if the queue is empty or not
template <typename T>
bool Queue<T>::isEmpty() {
    return head == nullptr;
}

// Define acceptable data types that the Queue can accept for the template
template class Queue<std::string>;
template class Queue<char>;
```

queue.h

```
#pragma once

#include "node.h"

template <typename T>
class Queue {
private:
    // Head points to the front of the queue
    Node<T>* head;
public:
    // We need a constructor and destructor
    Queue();
    ~Queue();

    // Enqueue adds a new element to the queue
    void enqueue(T newData);

    // Dequeue removes the front element from the queue
    T dequeue();

    // isEmpty checks to see if the queue is empty
    bool isEmpty();
};
```

5.4 Main Program

main.cpp

```
#include <iostream>
#include <string>

#include "node.h"
#include "stack.h"
#include "queue.h"
#include "fileUtil.h"
#include "util.h"

// Function to test the Node class
void testNode() {
    // Create the nodes on the stack, so we do not have to delete later
    Node<std::string> n1("node 1");
    Node<std::string> n2("node 2");
    Node<std::string> n3("node 3");

    // Set up the links
    n1.next = &n2;
    n2.next = &n3;

```

```cpp
    // Print out the data of each node in the linked list
    Node<std::string>* cur = &n1;
    while (cur != nullptr) {
        std::cout << cur->data << std::endl;
        cur = cur->next;
    }
}

// Function to test the Stack class
void testStack() {
    // Create a stack and add some data to it
    Stack<char> stack;
    stack.push('h');
    stack.push('s');
    stack.push('o');
    stack.push('J');

    // Print out the letters as we remove them from the stack
    while (!stack.isEmpty()) {
        std::cout << stack.pop();
    }
    std::cout << std::endl;

    try {
        // This should throw an error
        stack.pop();
    } catch (const std::invalid_argument& e) {
        std::cerr << e.what() << std::endl;
    }
}

// Function to test the Queue class
void testQueue() {
    // Create a queue and add some data to it
    Queue<char> queue;
    queue.enqueue('J');
    queue.enqueue('o');
    queue.enqueue('s');
    queue.enqueue('h');

    // Print out the letters as we remove them from the queue
    while (!queue.isEmpty()) {
        std::cout << queue.dequeue();
    }
    std::cout << std::endl;

    try {
        // This should throw an error
        queue.dequeue();
    } catch (const std::invalid_argument& e) {
        std::cerr << e.what() << std::endl;
    }
}

// Function to check if a string is a palindrome, minus whitespace and capitalization
bool isPalindrome(std::string word) {
    // Initialize an empty stack and queue for the checks
    Stack<char> wordStack;
    Queue<char> wordQueue;

    // Iterate through each character in the word to populate the stack and queue
    for (int i = 0; i < word.length(); i++) {
        char character = word[i];
        if (character == ' ') {
            // Go to next character because we are ignoring whitespace
            continue;
        } else if (character >= 'a' && character <= 'z') {
            // Adjust the character to make it uppercase by taking the difference between
            // the start of the lowercase letters and the start of the uppercase letters
            character -= 'a' - 'A';
        }
        // Add the character to both the stack and the queue
        wordStack.push(character);
        wordQueue.enqueue(character);
    }

    while (!wordStack.isEmpty() && !wordQueue.isEmpty()) {
        // Get the character from the top of the stack and queue
        char charFromStack = wordStack.pop();
        char charFromQueue = wordQueue.dequeue();

        if (charFromStack != charFromQueue) {
            // We can return false because we already know that
            // the string is not a palindrome
            return false;
        }
    }

    // The string is a palindrome
    return true;
}

int main() {
    std::cout << "----- Testing Node class -----" << std::endl;
    testNode();
    std::cout << std::endl;

    std::cout << "----- Testing Stack class -----" << std::endl;
    testStack();
    std::cout << std::endl;

    std::cout << "----- Testing Queue class -----" << std::endl;
    testQueue();
    std::cout << std::endl;

    std::cout << "----- Testing isPalindrome -----" << std::endl;
    std::cout << isPalindrome("racecar") << std::endl; // 1
    std::cout << isPalindrome("RaCecAr") << std::endl; // 1
    std::cout << isPalindrome("ra   c e   car") << std::endl; // 1
    std::cout << isPalindrome("4") << std::endl; // 1
    std::cout << isPalindrome("") << std::endl; // 1
    std::cout << isPalindrome("ABC") << std::endl; // 0
    std::cout << std::endl;

    std::cout << "----- Magic Items -----" << std::endl;
    try {
        // Read the file and store it in an array
        StringArr* data = readFile("magicitems.txt");

        // Only print out the palindromes
        for (int i = 0; i < data->length; i++) {
            if (isPalindrome(data->arr[i])) {
                std::cout << data->arr[i] << std::endl;
            }
        }

        // Clean up memory
        delete data;
    } catch (const std::invalid_argument& e) {
        std::cerr << e.what() << std::endl;
    }

    return 0;
}
```