CE221 Programming in C++
Part 7
The list and deque classes, Container Adapters, Associative Containners
The list Class 1

The **list** container is a sequence container that stores the sequence as dynamically-allocated cells each containing a pointer to a data item and pointers to the previous and next in the list. This means that no shifting is necessary when adding items to the beginning or middle or deleting them. However accessing items non-sequentially is not efficient so no subscripting or **at** functions are provided.

The space overheads may be significant if the items in the list do not use much memory – a linked list of numbers will occupy almost four times as much memory as a vector of the same capacity.

Iterators for list objects are bidirectional – they support **++** and **--**, but not **+** and **+=**.
The list Class 2

In addition to `push_back` the `list` class has a `push_front` member function for insertion at the front.

There is an `insert` method that takes an iterator as its first argument; the data item supplied as the second argument will be inserted in front of the position referenced by the iterator.

```cpp
#include <list>
list<int> l;
l.push_back(3);
l.push_front(4);
l.insert(++l.begin(), 44);
for_each(l.begin(), l.end(), print);
// use print function from part 6 slide 50
// output will be 4 44 3
```
The list Class 3

The no-argument `list` constructor initialises the list to be empty. There is also a two-argument constructor – the first argument is a repetition count and the second a value: `list<int> l(3, 4)` will initialise `l` to hold 3 copies of the value 4. (The second argument may be omitted in which case the value 0, or equivalent, will be used).

The first or last item of a list may be accessed using `front` or `back` and removed using `pop_front` or `pop_back`. Note that the effect of these functions when a list is empty is unpredictable. There is no function to remove items at other specified locations; to do this we need to use the iterator `erase` function.
All occurrences of a particular value may be removed using the \texttt{remove} function

```cpp
list<int> l(4, 5); // 5 5 5 5
l.insert(++l.begin(), 7); // 5 7 5 5 5
list<int>iterator pos = find(l.begin(), l.end(), 7);
l.insert(pos, 8); // 5 8 7 5 5 5
l.push_front(7); // 7 5 8 7 5 5 5
l.remove(7); // 5 8 5 5 5
cout << l.back(); // 5 will be output
l.pop_front(); // 8 5 5 5
```
The list Class 5

There is also a `remove_if` function that removes all values that satisfy a condition; the argument is a pointer to a function that takes a data item as its argument and returns a boolean value indicating whether that item satisfies the condition.

To remove all upper-case letters from a list `l1` of type `list<char>` we could use `l1.remove_if(isupper)`.

The following code will remove all negative values from the list `l2` of type `list<int>`.

```cpp
bool isneg(int i) {
    return i<0;
}

l2.remove_if(isneg);
```
The list Class 6

We cannot use the sort algorithm to sort lists, since they do not support random-access iterators. Hence there are two `sort` member functions to sort a list into ascending order; a no-argument version that sorts using `<` and a version with one argument similar to the last argument of the three-argument iterator `sort` function.

The function `unique` will remove from the list all but one of any sequence of adjacent equal items. If we want to eliminate all duplicate items from a list we need to sort it first so that equal items are adjacent. (Items are tested for equality using `==` so if the items in the list are objects of a class that class must have an `==` operator.)
The `splice` function can be used to move all of the elements of one list into another. `l1.splice(pos, l2)` moves all of the elements of `l2` into `l1` immediately in front of the element referenced by the iterator `pos`. After this `l2` will be empty. Note that no copying of elements takes place – only the pointers in the list cells are changed. The function can take extra arguments to allow part of `l2` to be moved.

The `merge` function can be used to move all of the elements of a sorted list into another, with the result being sorted. `s1.merge(s2)` moves all of the elements of `s2` into `s1` in appropriate positions to maintain correct ordering. If either of the lists is not sorted the order of elements is unpredictable. The functions can take an extra comparator argument to specify how to perform the ordering.
The deque Class

The deque container (pronounced as "deck" and being an abbreviation of double-ended queue) is similar to vector but supports efficient addition and removal at both ends. In order to achieve this it is not possible for subscripting to be as efficient as for vectors, but it is still relatively efficient.

The deque class provides all of the functionality of vector and additionally push_front, pop_front, front and back functions.

Programs that use this container need to include the header file <deque>.
The stack and queue Adaptors

The **stack** and **queue** container adaptors provide wrappers around sequence containers to allow them to be used as stacks and queues. A stack is a first-in-last-out sequence; items can be added to and removed from the top only and the only item that can be accessed directly is the top item. On the other hand a queue is a first-in-first-out sequence; items are added at the back but removed from the front and the only items that can be accessed directly are the front and back items.

A program that uses one of these adaptors requires the use of

```cpp
#include <stack>
```
or

```cpp
#include <queue>
```

1 In data structures courses it is normally assumed that the only item that can be accessed directly is the front one, but the C++ adaptor also allows access to the back one.
The stack Class 1

The `stack` class by default is implemented using a deque, but we may if we wish obtain a stack that is implemented using a vector or a list:

```cpp
stack<int> s1;               // uses deque<int>
stack<int, vector<int>> s2;  // uses vector<int>
```

We may initialise a stack to contain the contents of an existing sequence object of the same type as the class being used to implement the stack; the first item in the sequence will be at the bottom of the stack:

```cpp
list<int> l;
........
stack<int, list<int>> s3(l);
```
The stack Class 2

The main member functions of the **stack** class are **push**, **pop** and **top**. These are implemented using the functions **push_back**, **pop_back**, and **back**. **pop** and **top** have undefined behaviour if the stack is empty. In addition member functions **empty** and **size** (as described in part 6) are provided.

```cpp
include<stack>
stack<int> st;
st.push(5);
st.push(4);
while (!st.empty())
{
    cout << st.top() << endl;
    st.pop();
}
```
The queue Class 1

The `queue` class by default also uses a deque, but we may if we wish obtain a queue that is implemented using a list (but not a vector since the `vector` class does not support `pop_front`):

```cpp
#include<queue>

queue<char> q1;       // uses deque<char>
queue<char, list<char>> q2;  // uses list<char>
```

As with stacks we may initialise a queue to contain of the contents of an existing sequence object of the appropriate type; the first item in the sequence will be at the front of the queue.
The queue Class 2

The main member functions of the queue class are push, pop, front and back. The last three have undefined behaviour if the queue is empty. The push function adds to the back and the pop function removes the front element. Once again member functions empty and size are provided.

```cpp
queue<char> q;
q.push('x');
q.push('y');
cout << "back is " << q.back() << endl; // y
while (!q.empty())
{ cout << q.front() << endl;
  q.pop();
}                 // x will be output, then y
```
A *priority queue* is like a queue but items have priorities and when an item is added to the back it will advance ahead of any items with lower priority. The order in which items with equal priority will reach the head of the queue not defined. The precise state of the queue at any stage is also not defined, but it is always guaranteed that the head item will not have a lower priority than any other items in the queue.

The implementation of a priority queue normally uses a data structure called a heap; this may be efficiently implemented using any sequence with random-access iterators.

The *priority_queue* class by default uses a vector, but we may obtain a queue that is implemented using a deque.
The priority_queue Class 2

Programs that use the priority_queue adaptor need to use #include <queue>.

Priorities are by default compared using the < operator; in most applications the items in the queue will be class objects with the priority being a member of the class; we would need to write an operator< function for the class such that a<b returns true if a has lower priority than b.

If the item class already has a < operator that compares something other than the priorities, we need to supply a comparator class; this class should be specified in the declaration of the queue, as a third template argument e.g.

    priority_queue<Job, deque<Job>, MyComp> pq;
The comparator class needs to be a class with an `operator()` function that takes as arguments two data items `a` and `b` (or constant references to such items) and returns the value of `a<b` under the ordering to be used. For example if we wish to use a priority queue of objects of type `Job`, where `Job` is a class with a priority stored in a public instance variable `pri`, but has a `<` operator that does not compare priorities, we could write a class of the form

```cpp
class MyComp
{    bool operator()(const Job &a, const Job &b)
    { return a<b;
    }
}
```
The priority_queue Class 4

The member functions of the `priority_queue` class have the same names as those of the `stack` class.
In the following example we assume that the `YellowPerson` class has a priority that is set using the last argument of its constructor, the `<` operator compares priorities and the `<<` operator has been overloaded for the class.

```cpp
priority_queue<YellowPerson> pq;
pq.push(YellowPerson("homer", ……, 2));
pq.push(YellowPerson("marge", ……, 3));
cout << pq.top(); // will be marge
pq.pop();
pq.push(YellowPerson("bart", ……, 2));
cout << pq.top(); // could be homer or bart
```
C++ provides a template class called `pair` for storing pairs of objects. e.g.

```cpp
pair<int, bool> p(3, false);
```

We can access the elements of such a pair using the public data members `first` and `second`:

```cpp
cout << p.first << ' ' << p.second << endl;
```

If we wanted to supply a `pair` object as an argument to a function it would be inconvenient to specify its type explicitly so a template function called `make_pair` is provided

```cpp
myfun(make_pair("bart", 45));
```
Associative Containers 1

An associative container is used to store non-sequential data. The STL provides four such containers: set, multiset, map and multimap.

Although there is no concept of being sorted for the contents of an associative container, the STL classes store the items in order to allow efficient searching so the < operator must be defined for any objects in an associative container or a pointer to a comparison function must be supplied as an argument in the call to the constructor.

In addition to the member functions listed in part 6 that are common to all containers, the associative containers all have member functions called find, count, insert, lower_bound and upper_bound.
Associative Containers 2

Associative containers have bi-directional iterators but not random-access iterators.

The `find` method performs the same task as the `find` algorithm (but in a more efficient way), i.e. `c.find(x)` will give the same result as `find(c.begin(), c.end(), x)`. In the case of a multi-map or multi-set the iterator that is returned will reference the first occurrence in the underlying implementation; we can increment the iterator to reach other occurrences.

The `count` method returns the number of occurrences of its argument; for a set or map this will always be 0 or 1.
The `insert` method will insert an element into the collection; in the case of a set or map it will leave the collection unchanged if the element is already present. This method returns a result of type `pair<C::iterator, bool>` (where `C` is the instantiated container class, e.g. `set<int>`). The second element indicates whether anything was inserted; the first element is an iterator object that refers to the inserted element or the existing element with the same value if the element was not inserted.

```cpp
set<int> s;
......
if (s.insert(7).second)
    cout << "7 inserted";
else cout << "7 already present";
```
The `lower_bound` and `upper_bound` methods can be used to obtain start and end iterators for traversing a range of elements from a container.

`s.lower_bound(x)` will return an iterator referencing the first occurrence of an element greater than or equal to `x` whereas `s.upper_bound(y)` will return an iterator referencing the first occurrence of an element greater than `y`. Both will return an iterator equal to `s.end()` if no such element exists. To visit all of the elements in the range `x` to `y` inclusive we can use a loop of the form

```cpp
C::iterator end = s.upper_bound(y);
for (C::iterator it = s.lower_bound(x); it != end; it++)
    // process *it
```
The set Class 1

The `set` container is used to store sets – a set may not contain duplicate elements.

It has a no-argument constructor that will initialise a set to be empty, a copy constructor and a two-argument constructor that will initialise the contents of the set using all or part of an array or container object. The arguments are either start and end iterators or pointers as used on slide 15 of part 6. (If the array or container object contains duplicate values only one occurrence of each value will be stored in the set.)

The `<set>` header file needs to be included in programs that use sets (or multi-sets).
#include <set>
#include <string>

string a[6] = { "a", "b", "c", "d", "e", "f" };
vector<string> ( { "g", "h", "e", "i", "c", "a", "*****", "+++" } ) v;

set<string> A(a, a+6); // load A from a
set<string> B(v.begin(), v.begin()+6);
    // load B from first 6 elements of b
set<string> C;          // initially empty
Inserting an element into a set does not invalidate any iterators that already refer to elements of the set and removing an element using the `erase` method does not invalidate any iterators that do not refer to the deleted element.

Hence we could use the following code to remove all numbers greater than \( n \) from the set \( s \) (of type `set<int>`).

```cpp
set<int>::iterator it = s.begin(), it2;
while (it != s.end())
{
    it2 = it++;
    // need to advance it before deleting element
    if (*it2 > n)
        s.erase(it2);
}
```
The intersection of two sets $s_1$ and $s_2$ is a set containing all elements that occur in both $s_1$ and $s_2$, whereas the union of $s_1$ and $s_2$ is a set containing all elements that occur in $s_1$ or $s_2$ (or both). There are STL algorithms called `set_intersection` and `set_union` to obtain the intersection and union of two sets.

The first four arguments for each of these functions are iterators referencing the beginning and end of the two sets; a fifth argument is used to specify a set in which the result should be stored. This takes the form of a special sort of iterator called an inserter.

A function call of the form `inserter(s, s.begin())` is used to generate an inserter for a set $s$. (This function is declared in the `<iterator>` header file.)
The set Class 5

In the following code we assume that the sets A, B and C are as declared on slide 25.

```cpp
set_intersection(A.begin(), A.end(),
    B.begin(), B.end(),
    inserter(C, C.begin()));
// C contains "a", "c", "e"
C.clear();
set_union(A.begin(), A.end(),
    B.begin(), B.end(),
    inserter(C, C.begin()));
// C contains "a", "b", "c", "d", "e", "f",
// "g", "h", "i",
```
The `set` Class 6

There is also an algorithm called `set_difference`; this is used to obtain a set containing all of the elements that appear in the set specified by the first two arguments but do not appear in the set specified by the third and fourth arguments.

```cpp
C.clear();
set_difference(A.begin(), A.end(),
              B.begin(), B.end(),
              inserter(C, C.begin()));
// C contains "b", "d", "f"
C.clear();
set_difference(B.begin(), B.end(),
              A.begin(), A.end(),
              inserter(C, C.begin()));
// C contains "g", "i", "h"
```
The multiset Class

The multiset class is used for collections of elements which may contain duplicates; its methods are the same as those for set but some behave slightly differently. For example, the insert function will always succeed so the second element of the pair returned by the function will always be true.

The number of occurrences of an item in the union of two multisets is the larger of its occurrence counts from the two sets whereas in the intersection it is smaller of the two occurrence counts. An element will appear in the difference of two multisets if it occurs more times in the first set than in the second set; for example an item that occurs 5 times in the first set and 3 times in the second set will occur twice in the difference.
The map Class 1

A map is a collection of keys with associated values, where each key must be unique (but values may occur many times).

The map container in C++ is similar to Python's dict type – maps can be viewed as associative arrays and the values associated with keys can be accessed using subscripting.

Since the keys and values are normally of different types the map template class takes two type parameters, e.g.

```cpp
map<string, int> age;
```

Programs that use this class (or multimap) need to include the header file `<map>`.
The map Class 2

The **insert** method for the **map** class takes a pair as its argument – it is possible to write code such as

```cpp
pair<string, int> p("lisa", 14);
age.insert(p)
```

or

```cpp
age.insert(pair<string, int>("bart", 15));
```

but it is inconvenient to have to explicitly specify the type parameters. A more concise approach is to use the template function **make_pair** described on slide 18:

```cpp
age.insert(make_pair("homer", 55));
```

The **insert** method will not insert a pair if the key is already present in the map; it cannot be used to update the value associated with an existing key.
The map Class 3

An iterator for a map will reference a pair so we could print the contents of the `age` map using

```cpp
map<string, int>::const_iterator it;
for (it = age.begin(); it != age.end(); it++)
    cout << it->first << " : " << it->second << endl;
```

We could also use the `for_each` algorithm, supplying a function that takes a pair as an argument:

```cpp
void printPair(const pair<string, int> &p)
{ cout << p.first << " : " << p.second << endl; }
for_each(age.begin(), age.end(), printPair)
```
The \texttt{find} and \texttt{count} methods for maps take a key as an argument:

\begin{verbatim}
map<string,int>::iterator it = age.find("bart");
if (it != age.end())
    cout << "Bart's age is " << it->second << endl;
if (age.count("lisa") == 1)
    cout << "Lisa is " << age.find("lisa")->second;
\end{verbatim}
As in Python we can access the value associated with a key using subscripting:

```java
int bartsAge = age["bart"];  
```

If the key is not present in the map the value of the subscript operation will be 0 (or equivalent) – no exception will be thrown.

Subscripting is also used to update the value associated with a key:

```java
age["maggie"] = 1;
age["bart"];++;
```

If the key is not present in the map a new pair will be added.
In a *multi-map* a key may appear more than once. The methods for the `multimap` class are the same as those for `map`, apart from the absence of the subscripting operator, which cannot be supported since a key does not necessarily have a unique value.

To retrieve the value(s) associated with a key we have to use the `find` method. Since this returns a reference to the first occurrence of the key we can increment the iterator to find subsequent occurrences.
The `multimap` Class 2

Assuming `tel` is a multi-map for the storage of telephone numbers we could use the following code to print all of Bart's telephone numbers:

```cpp
multimap<string, int>::iterator it =
    tel.find("bart");

while (it != tel.end() && it->first == "bart")
{
    cout << it->second << endl;
    it++;
}
```

Note that in the underlying implementation only the keys are sorted so we cannot determine in which order the values will be printed.
To update a value associated with a key we must locate the pair using an iterator and then change the second element of the pair. Assuming one of Bart's numbers needs to be changed from 1234 to 5678 we might use

```cpp
multimap<string,int>::iterator it = tel.find("bart");
while (it != tel.end() && it->first == "bart")
{
  if (it->second == 1234)
  {
    it->second = 5678;
    break;
  }
  it++;
}
```