Lesson07---vector

【本节目标】

- **1.vector的介绍及使用**
- **2.vector深度剖析及模拟实现**
- **3.本节作业**

1.vector的介绍及使用

1.1 vector的介绍

> vector的文档介绍
>
> 1. vector是表示可变大小数组的序列容器。
> 2. 就像数组一样，vector也采用的连续存储空间来存储元素。也就是意味着可以采用下标对vector的元素进行访问，和数组一样高效。但是又不像数组，它的大小是可以动态改变的，而且它的大小会被容器自动处理。
> 3. 本质讲，vector使用动态分配数组来存储它的元素。当新元素插入时候，这个数组需要被重新分配大小为了增加存储空间。其做法是，分配一个新的数组，然后将全部元素移到这个数组。就时间而言，这是一个相对代价高的任务，因为每当一个新的元素加入到容器的时候，vector并不会每次都重新分配大小。
> 4. vector分配空间策略：vector会分配一些额外的空间以适应可能的增长，因为存储空间比实际需要的存储空间更大。不同的库采用不同的策略权衡空间的使用和重新分配。但是无论如何，重新分配都应该是对数增长的间隔大小，以至于在末尾插入一个元素的时候是在常数时间的复杂度完成的。
> 5. 因此，vector占用了更多的存储空间，为了获得管理存储空间的能力，并且以一种有效的方式动态增长。
> 6. 与其它动态序列容器相比（deques, lists and forward_lists）， vector在访问元素的时候更加高效，在末尾添加和删除元素相对高效。对于其它不在末尾的删除和插入操作，效率更低。比起lists和forward_lists统一的迭代器和引用更好。
>
> 学习方法：使用STL的三个境界：能用，明理，能扩展 ，那么下面学习vector，我们也是按照这个方法去学习

1.2 vector的使用

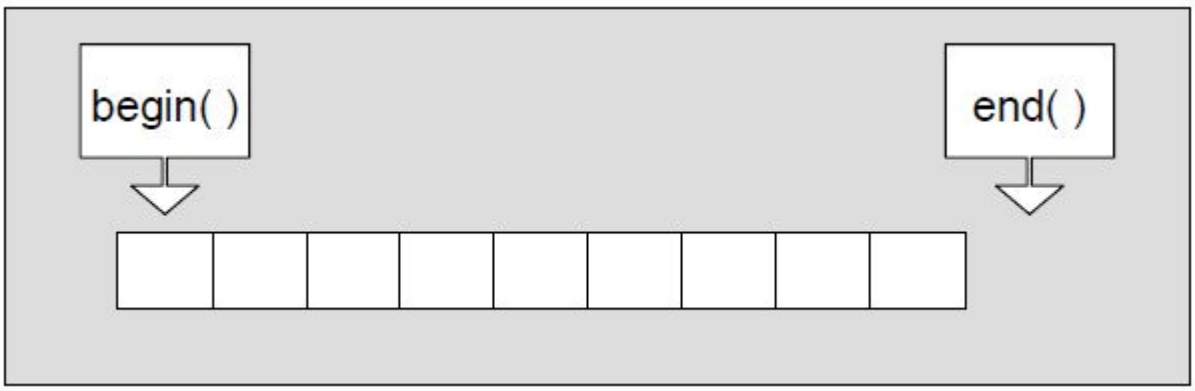
> vector学习时一定要学会查看文档：vector的文档介绍，vector在实际中非常的重要，在实际中我们熟悉常见的接口就可以，下面列出了**哪些接口是要重点掌握的。**
>
> #### 1.2.1 vector的定义

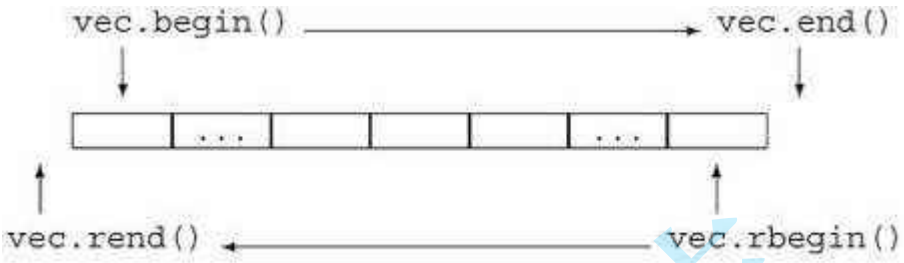
(constructor)构造函数声明	接口说明
vector()（重点）	无参构造
vector（size_type n, const value_type& val = value_type()）	构造并初始化n个val
vector (const vector& x);（重点）	拷贝构造
vector (InputIterator first, InputIterator last);	使用迭代器进行初始化构造

```
// constructing vectors
#include <iostream>
#include <vector>

int main ()
{
  // constructors used in the same order as described above:
  std::vector<int> first;                                // empty vector of ints
  std::vector<int> second (4,100);                       // four ints with value 100
  std::vector<int> third (second.begin(),second.end());  // iterating through second
  std::vector<int> fourth (third);                       // a copy of third

  // 下面涉及迭代器初始化的部分，我们学习完迭代器再来看这部分
  // the iterator constructor can also be used to construct from arrays:
  int myints[] = {16,2,77,29};
  std::vector<int> fifth (myints, myints + sizeof(myints) / sizeof(int) );

  std::cout << "The contents of fifth are:";
  for (std::vector<int>::iterator it = fifth.begin(); it != fifth.end(); ++it)
    std::cout << ' ' << *it;
  std::cout << '\n';

  return 0;
}
```

1.2.2 vector iterator 的使用

iterator的使用	接口说明
begin + end（重点）	获取第一个数据位置的iterator/const_iterator， 获取最后一个数据的下一个位置的iterator/const_iterator
rbegin + rend	获取最后一个数据位置的reverse_iterator，获取第一个数据前一个位置的reverse_iterator






```cpp
#include <iostream>
#include <vector>
using namespace std;

void PrintVector(const vector<int>& v)
{
    // const对象使用const迭代器进行遍历打印
    vector<int>::const_iterator it = v.begin();
    while (it != v.end())
    {
        cout << *it << " ";
        ++it;
    }
    cout << endl;
}

int main()
{
    // 使用push_back插入4个数据
    vector<int> v;
    v.push_back(1);
    v.push_back(2);
    v.push_back(3);
    v.push_back(4);

    // 使用迭代器进行遍历打印
    vector<int>::iterator it = v.begin();
    while (it != v.end())
    {
        cout << *it << " ";
        ++it;

    }
```

```
    cout << endl;

    // 使用迭代器进行修改
    it = v.begin();
    while (it != v.end())
    {
        *it *= 2;
        ++it;
    }

    // 使用反向迭代器进行遍历再打印
    vector<int>::reverse_iterator rit = v.rbegin();
    while (rit != v.rend())
    {
        cout << *rit << " ";
        ++rit;
    }
    cout << endl;

    PrintVector(v);

    return 0;
}
```

1.2.3 vector 空间增长问题

容量空间	接口说明
size	获取数据个数
capacity	获取容量大小
empty	判断是否为空
resize（重点）	改变vector的size
reserve （重点）	改变vector放入capacity

- capacity的代码在vs和g++下分别运行会发现，**vs下capacity是按1.5倍增长的，g++是按2倍增长的。**这个问题经常会考察，不要固化的认为，顺序表增容都是2倍，具体增长多少是根据具体的需求定义的。vs是PJ版本STL，g++是SGI版本STL。
- reserve只负责开辟空间，如果确定知道需要用多少空间，reserve可以缓解vector增容的代价缺陷问题。
- resize在开空间的同时还会进行初始化，影响size。

```
// vector::capacity
#include <iostream>
#include <vector>

int main ()
{
```

```
  size_t sz;
  std::vector<int> foo;
  sz = foo.capacity();
  std::cout << "making foo grow:\n";
  for (int i=0; i<100; ++i) {
    foo.push_back(i);
    if (sz!=foo.capacity()) {
      sz = foo.capacity();
      std::cout << "capacity changed: " << sz << '\n';
    }
  }
}
```

vs：运行结果：

```
making foo grow:
capacity changed: 1
capacity changed: 2
capacity changed: 3
capacity changed: 4
capacity changed: 6
capacity changed: 9
capacity changed: 13
capacity changed: 19
capacity changed: 28
capacity changed: 42
capacity changed: 63
capacity changed: 94
capacity changed: 141
```

g++运行结果：

```
making foo grow:
capacity changed: 1
capacity changed: 2
capacity changed: 4
capacity changed: 8
capacity changed: 16
capacity changed: 32
capacity changed: 64
capacity changed: 128
```

```
// vector::reserve
#include <iostream>
#include <vector>

int main ()
{
  size_t sz;
  std::vector<int> foo;
  sz = foo.capacity();
  std::cout << "making foo grow:\n";
  for (int i=0; i<100; ++i) {

    foo.push_back(i);
```

```
    if (sz!=foo.capacity()) {
      sz = foo.capacity();
      std::cout << "capacity changed: " << sz << '\n';
    }
  }

  std::vector<int> bar;
  sz = bar.capacity();
  bar.reserve(100);   // this is the only difference with foo above
  std::cout << "making bar grow:\n";
  for (int i=0; i<100; ++i) {
    bar.push_back(i);
    if (sz!=bar.capacity()) {
      sz = bar.capacity();
      std::cout << "capacity changed: " << sz << '\n';
    }
  }
  return 0;
}
```

```
// vector::resize
#include <iostream>
#include <vector>

int main ()
{
  std::vector<int> myvector;

  // set some initial content:
  for (int i=1;i<10;i++)
      myvector.push_back(i);

  myvector.resize(5);
  myvector.resize(8,100);
  myvector.resize(12);

  std::cout << "myvector contains:";
  for (int i=0;i<myvector.size();i++)
    std::cout << ' ' << myvector[i];
  std::cout << '\n';

  return 0;
}
```

1.2.3 vector 增删查改

vector增删查改	接口说明
push_back（重点）	尾插
pop_back （重点）	尾删
find	查找。（注意这个是算法模块实现，不是vector的成员接口）
insert	在position之前插入val
erase	删除position位置的数据
swap	交换两个vector的数据空间
operator[] （重点）	像数组一样访问

```
// push_back/pop_back
#include <iostream>
#include <vector>
using namespace std;

int main()
{
    int a[] = { 1, 2, 3, 4 };
    vector<int> v(a, a+sizeof(a)/sizeof(int));

    vector<int>::iterator it = v.begin();
    while (it != v.end()) {
        cout << *it << " ";
        ++it;
    }
    cout << endl;

    v.pop_back();
    v.pop_back();

    it = v.begin();
    while (it != v.end()) {
        cout << *it << " ";
        ++it;
    }
    cout << endl;

    return 0;
}
```

```
// find / insert / erase
#include <iostream>
#include <algorithm>

#include <vector>
```

```cpp
using namespace std;

int main()
{
    int a[] = { 1, 2, 3, 4 };
    vector<int> v(a, a + sizeof(a) / sizeof(int));

    // 使用find查找3所在位置的iterator
    vector<int>::iterator pos = find(v.begin(), v.end(), 3);

    // 在pos位置之前插入30
    v.insert(pos, 30);

    vector<int>::iterator it = v.begin();
    while (it != v.end()) {
        cout << *it << " ";
        ++it;
    }
    cout << endl;

    pos = find(v.begin(), v.end(), 3);
    // 删除pos位置的数据
    v.erase(pos);

    it = v.begin();
    while (it != v.end()) {
        cout << *it << " ";
        ++it;
    }
    cout << endl;

    return 0;
}
```

```cpp
// operator[]+index 和 C++11中vector的新式for+auto的遍历
// vector使用这两种遍历方式是比较便捷的。
#include <iostream>
#include <vector>
using namespace std;

int main()
{
    int a[] = { 1, 2, 3, 4 };
    vector<int> v(a, a + sizeof(a) / sizeof(int));

    // 通过[]读写第0个位置。
    v[0] = 10;
    cout << v[0] << endl;

    // 通过[i]的方式遍历vector
    for (size_t i = 0; i < v.size(); ++i)

        cout << v[i] << " ";
```

```
    cout << endl;

    vector<int> swapv;
    swapv.swap(v);

    cout << "v data:";
    for (size_t i = 0; i < v.size(); ++i)
        cout << v[i] << " ";
    cout << endl;

    cout << "swapv data:";
    for (size_t i = 0; i < swapv.size(); ++i)
        cout << swapv[i] << " ";
    cout << endl;

    // C++11支持的新式范围for遍历
    for(auto x : v)
        cout<< x << " ";
    cout<<endl;

    return 0;
}
```

1.2.4 vector 迭代器失效问题。（重点）

```
// insert/erase导致的迭代器失效
#include <iostream>
#include <algorithm>
#include <vector>
using namespace std;

int main()
{
    int a[] = { 1, 2, 3, 4 };
    vector<int> v(a, a + sizeof(a) / sizeof(int));

    // 使用find查找3所在位置的iterator
    vector<int>::iterator pos = find(v.begin(), v.end(), 3);

    // 删除pos位置的数据，导致pos迭代器失效。
    v.erase(pos);
    cout << *pos << endl; // 此处会导致非法访问

    // 在pos位置插入数据，导致pos迭代器失效。
    // insert会导致迭代器失效，是因为insert可
    // 能会导致增容，增容后pos还指向原来的空间，而原来的空间已经释放了。
    pos = find(v.begin(), v.end(), 3);
    v.insert(pos, 30);
    cout << *pos << endl; // 此处会导致非法访问

    return 0;

}
```

```
// 常见的迭代器失效的场景
#include <iostream>
#include <algorithm>
#include <vector>
using namespace std;

int main()
{
    int a[] = { 1, 2, 3, 4 };
    vector<int> v(a, a + sizeof(a) / sizeof(int));

    // 实现删除v中的所有偶数
    // 下面的程序会崩溃掉，如果是偶数，erase导致it失效
    // 对失效的迭代器进行++it，会导致程序崩溃
    vector<int>::iterator it = v.begin();
    while (it != v.end())
    {
        if (*it % 2 == 0)
            v.erase(it);

        ++it;
    }

    // 以上程序要改成下面这样，erase会返回删除位置的下一个位置
    vector<int>::iterator it = v.begin();
    while (it != v.end())
    {
        if (*it % 2 == 0)
            it = v.erase(it);
        else
            ++it;
    }

    return 0;
}
```


1.2.5 vector 在OJ中的使用。

1. 只出现一次的数字i

```
class Solution {
public:
    int singleNumber(vector<int>& nums) {
        int value = 0;
        for(size_t i =0 ; i < nums.size(); ++i){
            value ^= nums[i];
        }
        return value;
    }
};
```

2. 杨辉三角OJ

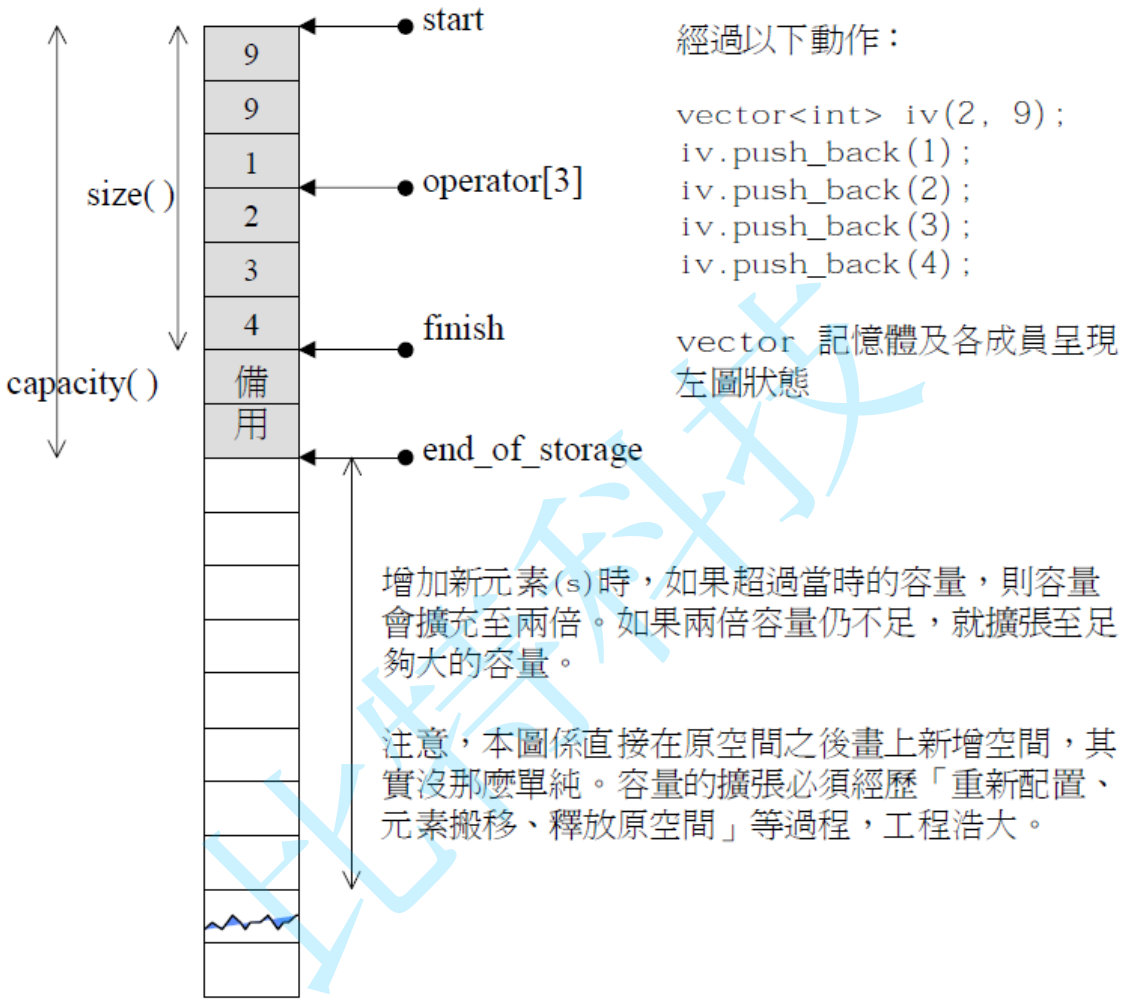
```
// 涉及resize / operator[]
class Solution {
public:
    // 核心思想：找出杨辉三角的规律，发现每一行头尾都是1，中间第[j]个数等于上一行[j-1]+[j]
    vector<vector<int>> generate(int numRows) {
        vector<vector<int>> vv;
       // 先开辟杨辉三角的空间
        vv.resize(numRows);
        for(size_t i = 1; i <= numRows; ++i)
        {
            vv[i-1].resize(i, 0);

            // 每一行的第一个和最后一个都是1
            vv[i-1][0] = 1;
            vv[i-1][i-1] = 1;
        }

        for(size_t i = 0; i < vv.size(); ++i)
        {
            for(size_t j = 0; j < vv[i].size(); ++j)
            {
                if(vv[i][j] == 0)
                {
                    vv[i][j] = vv[i-1][j-1] + vv[i-1][j];
                }
            }
        }

        return vv;
    }
};
```

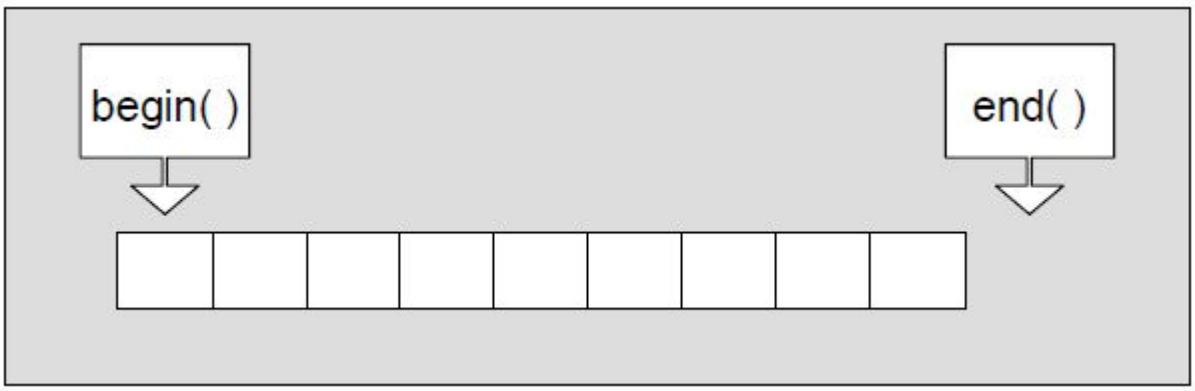
总结：通过上面的练习我们发现vector常用的接口更多是插入和遍历。遍历更喜欢用数组operator[i]的形式访问，因为这样便捷。课下除了再自己实现一遍上面课堂讲解的OJ练习，请自行完成下面题目的OJ练习。以此增强学习vector的使用。

3. 删除排序数组中的重复项 OJ

4. 只出现一次的数ii OJ
5. 只出现一次的数iii OJ
6. 数组中出现次数超过一半的数字 OJ
7. 电话号码字母组合OJ
8. 连续子数组的最大和 OJ

2.vector深度剖析及模拟实现





2.1 std::vector的核心框架接口的模拟实现bit::vector

```
#include <iostream>
```

```cpp
using namespace std;
#include <assert.h>

// 注意这里namespace大家下去就不要取名为bit了，否则出去容易翻车。^^
namespace bit
{
template<class T>
class vector
{
public:
    // Vector的迭代器是一个原生指针
    typedef T* iterator;
    typedef const T* const_iterator;

    iterator begin() { return _start; }
    iterator end() { return _finish; }

    const_iterator cbegin() const { return _start; }
    const_iterator cend() const { return _finish; }

        // construct and destroy
    vector()
        : _start(nullptr)
        , _finish(nullptr)
        , _endOfStorage(nullptr)
    {}

    vector(int n, const T& value = T())
        : _start(nullptr)
        , _finish(nullptr)
        , _endOfStorage(nullptr)
    {
        reserve(n);
        while (n--)
        {
            push_back(value);
        }
    }

    // 如果使用iterator做迭代器，会导致初始化的迭代器区间[first,last)只能是vector的迭代器
    // 重新声明迭代器，迭代器区间[first,last]可以是任意容器的迭代器
    template<class InputIterator>
    vector(InputIterator first, InputIterator last)
    {
        reserve(last - first);
        while (first != last)
        {
            push_back(*first);
            ++first;
        }
    }


    vector(const vector<T>& v)
```

```
        : _start(nullptr)
        , _finish(nullptr)
        , _endOfStorage(nullptr)
    {
        reserve(v.capacity());
        iterator it = begin();
        const_iterator vit = v.cbegin();
        while (vit != v.cend())
        {
            *it++ = *vit++;
        }

        _finish = _start + v.size();
        _endOfStorage = _start + v.capacity();
    }

    vector<T>& operator= (vector<T> v)
    {
        swap(v);
        return *this;
    }

    ~vector()
    {
        delete[] _start;
        _start = _finish = _endOfStorage = nullptr;
    }

        // capacity
         size_t size() const { return _finish - _start; }
    size_t capacity() const { return _endOfStorage - _start; }

    void reserve(size_t n)
    {
        if (n > capacity())
        {
            size_t oldSize = size();
            T* tmp = new T[n];

            // 这里直接使用memcpy是有问题的
            // 以后我们会用更好的方法解决
            //if (_start)
            //  memcpy(tmp, _start, sizeof(T)*size);
            if (_start)
            {
                for (size_t i = 0; i < oldSize; ++i)
                    tmp[i] = _start[i];
            }

            _start = tmp;
            _finish = _start + size;
            _endOfStorage = _start + n;

        }
```

```
    }

    void resize(size_t n, const T& value = T())
    {
        // 1.如果n小于当前的size，则数据个数缩小到n
        if (n <= size())
        {
            _finish = _start + n;
            return;
        }

        // 2.空间不够则增容
        if (n > capacity())
            reserve(n);

        // 3.将size扩大到n
        iterator it = _finish;
        iterator _finish = _start + n;
        while (it != _finish)
        {
            *it = value;
            ++it;
        }
    }

        ///////////////access//////////////////////////////
        T& operator[](size_t pos){return _start[pos];}
        const T& operator[](size_t pos)const {return _start[pos];}

        ///////////////modify//////////////////////////////
    void push_back(const T& x){insert(end(), x);}
    void pop_back(){erase(--end());}

        void swap(vector<T>& v)
    {
        swap(_start, v._start);
        swap(_finish, v._finish);
        swap(_endOfStorage, v._endOfStorage);
    }

    iterator insert(iterator pos, const T& x)
    {
        assert(pos <= _finish);

        // 空间不够先进行增容
        if (_finish == _endOfStorage)
        {
            size_t size = size();
            size_t newCapacity = (0 == capacity())? 1 : capacity() * 2;
            reserve(newCapacity);

            // 如果发生了增容，需要重置pos

            pos = _start + size;
```

```
        }

        iterator end = _finish - 1;
        while (end >= pos)
        {
            *(end + 1) = *end;
            --end;
        }

        *pos = x;
        ++_finish;
        return pos;
    }

    // 返回删除数据的下一个数据
    // 方便解决:一边遍历一边删除的迭代器失效问题
    iterator erase(Iterator pos)
    {
        // 挪动数据进行删除
        iterator begin = pos + 1;
        while (begin != _finish) {
            *(begin - 1) = *begin;
            ++begin;
        }

        --_finish;

        return pos;
    }
private:
    iterator _start;        // 指向数据块的开始
    iterator _finish;       // 指向有效数据的尾
    iterator _endOfStorage; // 指向存储容量的尾
};
}
```

2.2 对bit::vector核心接口的测试

```
// constructing vectors
void TestVector1()
{
    // constructors used in the same order as described above:
    bite::vector<int> first;                                // empty vector of ints
    bite::vector<int> second(4, 100);                       // four ints with value 100
    bite::vector<int> third(second.Begin(), second.End());  // iterating through second
    bite::vector<int> fourth(third);                        // a copy of third

    // the iterator constructor can also be used to construct from arrays:
    int myints[] = { 16, 2, 77, 29 };
    bit::vector<int> fifth(myints, myints + sizeof(myints) / sizeof(int));

    std::cout << "The contents of fifth are:";
    for (bit::vector<int>::iterator it = fifth.begin(); it != fifth.end(); ++it)
```

```
        std::cout <<*it<< " ";
    std::cout << endl;

    // 测试T是string时，拷贝问题
    bit::vector<string> strV;
    strV.PushBack("1111");
    strV.PushBack("2222");
    strV.PushBack("3333");
    strV.PushBack("4444");
    for (size_t i = 0; i < strV.size(); ++i)
    {
        cout << strV[i] << " ";
    }
    cout << endl;
}

//vector iterator的使用
void PrintVector(const bite::vector<int>& v)
{
    // 使用const迭代器进行遍历打印
    bit::vector<int>::const_iterator it = v.begin();
    while (it != v.end())
    {
        cout << *it << " ";
        ++it;
    }
    cout << endl;
}

void TestVector2()
{
    // 使用push_back插入4个数据
    bite::vector<int> v;
    v.push_back(1);
    v.push_back(2);
    v.push_back(3);
    v.push_back(4);
    PrintVector(v);

    // 使用迭代器进行修改
    auto it = v.begin();
    while (it != v.end())
    {
        *it *= 2;
        ++it;
    }

    PrintVector(v);

    // 这里可以看出C++11支持iterator及接口，就支持范围for
    for(auto e : v)
        cout<<e<<" ";

}
```

```cpp
// find / insert / erase
void TestVector3()
{
    int a[] = { 1, 2, 3, 4 };
    bite::vector<int> v(a, a + sizeof(a) / sizeof(a[0]));

    // 使用find查找3所在位置的iterator
    auto pos = find(v.begin(), v.end(), 3);

    // 在pos位置之前插入30
    v.insert(pos, 30);
    PrintVector(v);

    // 删除pos位置的数据
    pos = find(v.begin(), v.end(), 3);
    v.Erase(pos);
    PrintVector(v);
}

// iterator失效问题
void TestVector4()
{
    int a[] = { 1, 2, 3, 4 };
    bite::vector<int> v(a, a + sizeof(a) / sizeof(a[0]));

    // 删除pos位置的数据，导致pos迭代器失效
    auto pos = find(v.begin(), v.end(), 3);
    v.erase(pos);
    cout << *pos << endl; // 此处会导致非法访问

    // 在pos位置插入数据，导致pos迭代器失效。
    // insert会导致迭代器失效，是因为insert可
    // 能会导致增容，增容后pos还指向原来的空间，而原来的空间已经释放了。
    pos = find(v.begin(), v.end(), 3);
    v.insert(pos, 30);
    cout << *pos << endl; // 此处会导致非法访问

    // 实现删除v中的所有偶数
    // 下面的程序会崩溃掉，如果是偶数，erase导致it失效
    // 对失效的迭代器进行++it，会导致程序崩溃
    auto it = v.begin();
    while (it != v.end())
    {
        if (*it % 2 == 0)
            v.erase(it);

        ++it;
    }

    // 以上程序要改成下面这样，erase会返回删除位置的下一个位置
    it = v.begin();

    while (it != v.end())
```

```
    {
        if (*it % 2 == 0)
            it = v.erase(it);
        else
            ++it;
    }
}

/*
迭代器失效场景总结：
  1. 删除pos迭代器位置所指向元素没有及时给pos赋值，比如v.erase(pos)
  2. 可能会引起vector底层空间改变的操作，比如：push_back、insert、resize、reserve等
*/
```

3.本节作业

1. 练习vector的基本操作的使用
2. 实现本节课件中所有的OJ题目
3. 模拟实现vector及严格测试。
4. 本节内容很重要，希望大家认真完成。**做好第1 2点作业**，可以帮助你更好的**通过笔试。做好第3点**可以更好的帮助你**提高代码能力及通过面试。**