**Smart Pointer:**

*Smart pointers* are objects looks like built-in pointers (dumb pointer), but to offer greater functionality. They have a variety of applications, including resource management and the automation of repetitive coding tasks.

When we replace the *dumb* pointers with smart pointer then, gain control over the following aspects of pointer behaviours:

**Construction and destruction.** We can determine, what happens when a smart pointer is created and destroyed. Assign default value to the pointer (get rid from uninitialized pointers), automatically destroyed when goes out of scope (or ref count decrease when use shared ptr and destroyed when ref cnt=0) to eliminating resource leaks, lazy initialization (not initialized when created, until generate request for use) and many more.

**Copying and assignment.** We can decide the action, when a smart pointer is copied or is involved in an assignment like shallow copy or deep copy.

**Dereferencing.** What should happen when a client refers to the object pointed to by a smart pointer? For example, use smart pointers to help implement the lazy fetching strategy.

Most smart pointer templates look something like this:

template<class T> // template for smart pointer objects

class SmartPtr {

public:

SmartPtr(T\* realPtr = 0); // create a smart ptr to an obj given

// a dumb ptr to it; uninitialized

// ptrs default to 0 (null)

SmartPtr(const SmartPtr& rhs);// copy a smart ptr

~SmartPtr(); // destroy a smart ptr

// make an assignment to a smart ptr

SmartPtr& operator=(const SmartPtr& rhs);

T\* operator->() const; // dereference a smart ptr to get at

// a member of what it points to

T& operator\*() const; // dereference a smart ptr

private:

T \*pointee; // what the smart ptr points to

};

For smart pointer classes where copying and assignment are not allowed, both (copy constructor and assignment operator) would typically be declared as private.

The two dereferencing operators are declared const, because dereferencing a pointer doesn't modify it (though it may lead to modification of what the pointer points to).

Finally, each smart pointer containing a dumb pointer-to-T within it. It is this dumb pointer that does the actual pointing.

**Construction, Assignment and Destruction of smart pointer.**

Locate an object to point to (typically by using the smart pointer's constructor arguments), then make the smart pointer's internal dumb pointer point there. If no object can be located, set the internal pointer to 0 or signal an error (possibly by throwing an exception).

SmartPtr(T \*realPtr = 0) :pointee(realPtr) {}

How to implement the smart pointer's copy constructor, assignment operator(s) and

destructor.

If a smart pointer *owns* the object it points to, it is responsible for deleting that object when it (the smart pointer) is destroyed. The object pointed to by the smart pointer is dynamically allocated.

Consider the auto\_ptr template from the standard C++ library. An auto\_ptr object is a smart pointer that points to a heap-based object until it (the auto\_ptr) is destroyed. The auto\_ptr template might be implemented like:

template<class T>

class auto\_ptr {

public:

auto\_ptr(T \*ptr = 0): pointee(ptr) {}

~auto\_ptr() { delete pointee; }

auto\_ptr(auto\_ptr<T>& rhs); //copy constructor, non-const param

auto\_ptr<T>& operator=(auto\_ptr<T>& rhs); // assignment operator

// Also having non-const param.

//*Note:Usually copy ctor & assigenment operator expect const params.*

...

private:

T \*pointee;

};

But what should happen when an auto\_ptr is copied or assigned?

auto\_ptr<TreeNode> ptn1(new TreeNode);

auto\_ptr<TreeNode> ptn2 = ptn1; // call to copy ctor then, what should happen?

auto\_ptr<TreeNode> ptn3;

ptn3 = ptn2; // call to operator= then, what should happen?

Now we can do two things:

1. We just copied the internal dumb pointer and end up with two auto\_ptrs pointing to the same object (Problem: If one pointer deleted it will easily invalidate others).
2. An alternative would be to create a new copy of what was pointed to by calling new. That would guarantee we didn't have too many auto\_ptrs pointing to a single object

(Problem: Performance bottleneck, too many new object).

The problems would vanish if auto\_ptr prohibited copying and assignment, but a more flexible solution was adopted for the auto\_ptr classes: object ownership is *transferred* when an auto\_ptr is copied or assigned:

//Copy Constructor implementation

template<class T>

auto\_ptr<T>::auto\_ptr(auto\_ptr<T>& rhs)

{

pointee = rhs.pointee; // transfer ownership of \*pointee to \*this

rhs.pointee = 0; // rhs no longer owns anything

//If rhs parameter deleacred as const then rhs.pointee = 0 will //become invalid.(no ownership will transfer)

}

//Assignment operator Implementation.

template<class T>

auto\_ptr<T>& auto\_ptr<T>::operator=(auto\_ptr<T>& rhs)

{

if (this == &rhs) // do nothing if this

return \*this; // object is being assigned to itself

delete pointee; // delete currently owned object

pointee = rhs.pointee; // transfer ownership of

rhs.pointee = 0; // \*pointee from rhs to \*this

return \*this;

//If rhs parameter deleacred as const then rhs.pointee = 0 will

//become invalid. (no ownership will transfer)

}

Now, because of object ownership is transferred when auto\_ptr's copy constructor is called, passing auto\_ptrs by value is often a *very* bad idea. Here's why:

// This function will often lead to disaster.

void printTreeNode(ostream& s, auto\_ptr<TreeNode> p){

s << \*p;

}

int main()

{

auto\_ptr<TreeNode> ptn(new TreeNode);

...

printTreeNode(cout, ptn); // pass auto\_ptr by value

... // ownership transfer to p

} // now ptn is null.

When printTreeNode's parameter p is initialized (by calling auto\_ptr's copy constructor), ownership of the object pointed to by ptn is transferred to p. When printTreeNode finishes executing, p goes out of scope and its destructor deletes what it points to. Now ptn, no longer points to anything (pointing to null).

Instead of that, if we write a method like this: (Pass-by-reference-to-const)

void printTreeNode(ostream& s, const auto\_ptr<TreeNode> &p){

//No copy ctor call, for p.

s << \*p; //Since no new obj created, so no ownership transfer

}

***Note****: copy ctor & assignment operator functions normally take const parameter. But in*

*Case of auto\_ptr there is changes in parameters during the copy or the assignment. In other words, auto\_ptr objects are modified if they are copied or are the source of an assignment. So, it cannot be const. Let suppose, if we necessarily declared parameter as const, C++ offers const\_cast to remove const-ness from object.*

A smart pointer's destructor often looks like this:

template<class T>

SmartPtr<T>::~SmartPtr()

{

if (*\*this owns \*pointee*) {

delete pointee;

}

}

There is no need to check if condition for auto\_ptr, because auto\_ptr always owns what it points to. But for another smart pointer (like shared\_ptr) need to check total ref cnt before delete the pointee.

**Implementing the Dereferencing Operators:**

template<class T>

T& SmartPtr<T>::operator\*() const{

*perform "smart pointer" processing;*

return \*pointee;

}

**Similarly,**

T\* SmartPtr<T>::operator->() const{

*perform "smart pointer" processing;*

return pointee;

}

First the function does whatever processing is needed to initialize. For example, if lazy fetching is being used, the function performs required action to make Pointee valid. Once pointee is valid, the operator\* or-> function just returns a reference or pointer respectively to the pointed-to object.

***Note:*** *De-reference function return type is either pointer or reference not an object. pointee need not point to an object of type T; it may point to an object of a class derived from T. If that is the case and dereferencing function returns a T object instead of a pointer or reference to the actual derived class object, our function will return an object of the wrong type! (slicing problem).*

**Testing Smart Pointers for Null-ness:**

SmartPtr<TreeNode> ptn;

...

We test Null like below:

if (ptn == 0) ... // error!

if (ptn) ... // error!

if (!ptn) ... // error!

Error, because of required function is absent form SmartPtr class. One way is adding isNull member function to our smart pointer classes, to test the nullability. But we want to keep SmartPtr behavior like dumb pointer. How we can do that?

We can write the implicit conversion operator method which convert pointee type to void\*.

template<class T>

class SmartPtr {

public:

...

operator void\*(){ ,

if(pointee==NULL)

return NULL; // returns 0 if the smart ptr is null

else

return (void\*)1; // otherwise return nonzero

};

In below case every time it calls conversion function and return zero if null otherwise return non zero value.

if (ptn == 0) ... // fine

if (ptn) ... // fine

if (!ptn) ... // fine

But this approach has major disadvantage, Lets suppose we have a two smart pointer of different type:

SmartPtr<Apple> pa;

SmartPtr<Orange> po;

...

if (pa == po) ... // Even we don’t have operator == inside

// SmartPtr, this compile!

Why? Because both smart pointers can be implicitly converted into void\* pointers, and there is a built-in comparison function (compiler generated) which makes comparison.

We may think to overload the operator!, But problem will still remain same.

template<class T>

class SmartPtr {

public:

...

bool operator!() const; // returns true if and only

... // if the smart ptr is null

};

SmartPtr<TreeNode> ptn;

...

if (!ptn) { // fine

... // ptn is null

}

else {

... // ptn is not null

}

but below checks are failed:

if (ptn == 0) ... // still an error

if (ptn) ... // also an error

The only risk for mixed-type comparisons is statements such as these:

SmartPtr<Apple> pa;

SmartPtr<Orange> po;

...

if (!pa == !po) ... // this compiles and gives un-desirable result

We will discuss later, how to implement the null check for SmartPtr class.

**Converting Smart Pointers to Dumb Pointers:**

Consider the below class:

class Person {

int age;

char\* pName;

char\* initName(char \*nm) {

int size = strlen(nm);

char \*tmp = new char[size];

memcpy(tmp, nm, strlen(nm));

tmp[size] = '\0';

return tmp;

}

public:

Person() : pName(0), age(0) {}

Person(char\* nm, int age) : pName(initName(nm)), age(age){ }

~Person() {

delete pName;

age = 0;

}

void Display()const {

printf("Name = %s Age = %d \n", pName, age);

}

void Shout()const{

printf("Ooooooooooooooooo\n");

}

};

Bool validatePerson (const Pesron \*p) {

//Validate the person base on some logic return true if success otherwise false.

} //or both are same.

Bool validatePerson (const SmartPtr<Person> &p){

}

Now,

SmartPtr<Person> sp1(new Person(“Rajeev Sharma”,32));

validatePerson(sp1); //Error, cannot convert SmartPtr<Person> to Person\*

Yes, we can do it like this: validatePerson(&\*sp1); But it is look like ugly.

The call can be made to succeed by adding to the smart pointer-to-T template an implicit conversion operator to a dumb pointer-to-T:

template<class T> // as before

class SmartPtr {

public:

...

operator T\*() { return pointee; }

...

};

validatePerson(sp1); //Now this call will work. ***And apparently, this function also eliminates the problem of testing for null-ness:***

if (sp1 == 0) ... // fine, converts sp1 to a Person\*

if (sp1) ... // ditto

if (!sp1) ... // ditto (reprise)

However, it also has same some loopback. They make it easy for clients to program directly with dumb pointers bypassing the smart pointer object like:

Void UpdateID(SmartPtr<Person> &person){

Person \*p= person; //Converts SmartPtr<Person> to Person\*

…

//Now client do anything they want using row pointer \*p, bypassing SmartPtr.

//Even if parameter is const, then apply const\_cast to remove the const-ness

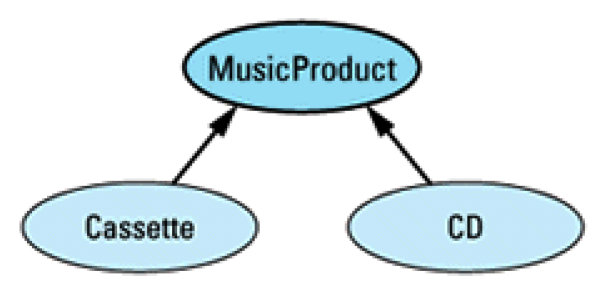
}

Usually, the "smart" behavior provided by a smart pointer is an essential component of our design, so allowing clients to use dumb pointers typically leads to disaster. For example, if SmartPtr implements the reference-counting strategy, allowing clients to manipulate dumb pointers directly will almost certainly lead to bookkeeping errors that corrupt the reference-counting data structures.

***Moral****: Don't provide implicit conversion operators to dumb pointers unless there is a compelling reason to do so. Provide isNull() method to check the nullability* ***(according to my observation).***

**Smart Pointers and Inheritance-Based Type Conversions:**

Suppose we have a public inheritance hierarchy modeling consumer product for storing music:



class **MusicProduct** {

public:

MusicProduct(const string& title);

virtual void play() const = 0;

virtual void displayTitle() const = 0;

...

};

class **Cassette**: public **MusicProduct** {

public:

Cassette(const string& title);

virtual void play() const;

virtual void displayTitle() const;

...

};

class **CD**: public **MusicProduct** {

public:

CD(const string& title);

virtual void play() const;

virtual void displayTitle() const;

...

};

Suppose we have a function that, given a MusicProduct object, displays the title of the product and then plays it:

void displayAndPlay(const MusicProduct\* pmp, int numTimes)

{

for (int i = 1; i <= numTimes; ++i) {

pmp->displayTitle();

pmp->play();

}

}

SmartPtr<Cassette> funMusic(new Cassette("Alapalooza"));

SmartPtr<CD> nightmareMusic(new CD("Disco Hits of the 70s"));

displayAndPlay(funMusic, 10); // error!

displayAndPlay(nightmareMusic, 0); // error!

We can remove above error if, we write a conversion method form smartPtr to dumb pointer.

But we have already seen, writing those conversion in SmartPtrs are not advisable.

look what happens if we replace the dumb pointers with their allegedly smart counterparts:

void displayAndPlay(const SmartPtr<MusicProduct>& pmp, int numTimes);

Still, they won't compile because there is no conversion from a SmartPtr<CD> or a SmartPtr<Cassette> to a SmartPtr<MusicProduct>

We may think, to provide an implicit type conversion operator for smart pointer class to which it should be implicitly convertible. Like:

template<class T> // as before

class SmartPtr {

public:

...

operator SmartPtr<T>() {

return SmartPtr<T>(pointee);

}

...

Private:

T \*pointee;

};

The above conversion method will say, pointee is converted to SmartPtr whatever it has its type. Suppose if pointee type is CD then conversion method converts it into SmartPtr<CD>.

displayAndPlay(funMusic, 10); // error!

displayAndPlay(nightmareMusic, 0); // error!

A reference of type "SmartPtr<MusicProduct> &" cannot be initialized with a value of type "SmartPtr<Cassette >" and “SmartPtr<CD>” respectively. It will Compile if we change the respective function like below:

displayAndPlay(const SmartPtr<Cassette>& pmp,int numTimes);//For Cassette

displayAndPlay(const SmartPtr<CD>& pmp, int numTimes); //For CD.

But this is not what we want. Now we changed our SmartPtr class like below:

template<class T> // template class for smart

class SmartPtr {

public:

SmartPtr(T\* realPtr = 0);

T\* operator->() const;

T& operator\*() const;

...

...

template<class newType> //template function for

operator SmartPtr<newType>() //implicit conversion

{

return SmartPtr<newType>(pointee);

}

Private:

T \*pointee

...

};

Suppose a compiler has a smart pointer-to-T object (Cassette or CD) and need to convert into a smart pointer-to-base-class-of-T(MusicProduct). The compiler checks the class definition for SmartPtr<T> to see, if the requisite conversion operator (below) is declared or not:

operator SmartPtr<MusicProduct>() {

return SmartPtr<MusicProduct>(pointee);

}

Whereas our T \*pointee is either CD or cassette. And obviously this method not found.

Then compiler checks to see if there's a member function template it can instantiate that would let it perform the conversion it's looking for. It finds such a template (the one taking the formal type parameter newType), so it instantiates the template with newType bound

to the base class of T that's the target of the conversion.

At that point, the only question is whether the code for the instantiated member function will compile. For it to compile, it must be legal to pass the (dumb) pointer pointee to the constructor for the smart pointer-to-base-of-T. pointee is of type T, so it is certainly legal

to convert it into a pointer to its (public or protected) base classes. Hence, the code for the type conversion operator will compile, and the implicit conversion from smart pointer-to-T to smart pointer-to-base-of-T will succeed.

**Smart Pointers and const:**

SmartPtr<CD> p; // non-const object, non-const pointer

SmartPtr<const CD> p; // const object, non-const pointer

const SmartPtr<CD> p = &goodCD; // non-const object, const pointer

const SmartPtr<const CD> p = &goodCD; // const object, const pointer

we can initialize const pointers with non-const pointers and we can initialize pointers to const objects with pointers to non-consts.

CD \*pCD = new CD("Famous Movie Themes");

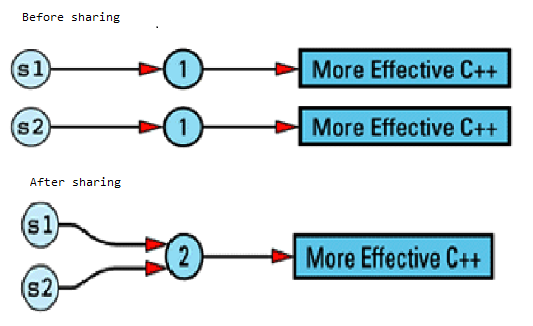
const CD \* pConstCD = pCD; // fine

But look what happens if we try the same thing with smart pointers:

SmartPtr<CD> pCD = new CD("Famous Movie Themes");

SmartPtr<const CD> pConstCD = pCD; // fine?

**Reference Counting. (Smart Reference Proxy)**



class RCString

{

private:

struct StringValue {

int refCount;

char \*data;

StringValue(const char \*initValue);

~StringValue();

};

StringValue \*value; // value of this String

public:

RCString(const char \*initValue = "");

RCString(const RCString& rhs);

RCString& operator=(const RCString& rhs);

char& operator[](int index)const; //For Const String

char& operator[](int index); //For non-Const String.

~RCString();

void showStr()const {

cout <<"Data:: "<<value->data<<", Address::"<< &(value->data)<<endl;

cout << "Ref cnt:: " << value->refCount<<endl;

cout << "---------------\n";

}

};

//Implementation of the const version of this function is straightforward, because it's a //read-only operation; the value of the string can't be affected

char& RCString::operator[](int index)const

{

if ((strlen(value->data) - 1) > index)

return value->data[index];

///else throw ArrayOutOfBoundExeception

}

//Non-const version, Since we don’t know whether it is read or write, so we are creating //new set of stringValue therefore it cannot create problem while write operation happening.

char& RCString::operator[](int index) //throw ArrayOutOfBoundExeception,

//Implement later..

{

// if we're sharing a value with other String objects, break off a

// separate copy of the value for ourselves

if (value->refCount > 0)

{

// decrement current value's refCount, because we won't be using

// that value any more

value->refCount--;

value = new StringValue(value->data); // make a copy of the value

// for ourselves.

}

// return a reference to a character inside our unshared StringValue

// object

return value->data[index];

}

RCString& RCString::operator=(const RCString& rhs)

{

if (this->value == rhs.value)

return \*this;

this->~RCString();

this->value = rhs.value;

this->value->refCount++;

return \*this;

}

RCString::~RCString()

{

if ((--value->refCount) == 0) {

value->~StringValue();

}

}

RCString::RCString(const char \*initValue): value(new StringValue(initValue))

{

}

RCString::RCString(const RCString& rhs):value(rhs.value) {

++value->refCount;

}

RCString::StringValue::StringValue(const char \*initValue):refCount(1) {

data = new char[strlen(initValue) + 1];

strcpy(data, initValue);

}

RCString::StringValue::~StringValue(){

delete[] data;

}

int main() {

Problem with Write operation-When we modify a String's value, we must be careful to avoid modifying the value of other String objects. Because we all share same value among classes and unfortunately, there is no way for C++ compilers to tell us whether a use of operator [] is for a read or a write.

//Write Operation

RCString str1("RajeevKumarSharma"); //str1->refCount=1

RCString str2(str1); //str1->refCount=2,str2->refCount=2

//Both have same copy of string.

str2[0] = 'X'; //Now we will be modifying str2, but It should not

//be impact str1.

str1.showStr(); //str1->refCount=1

str2.showStr(); //str2->refCount=1, It have separate copy of string //"XajeevKumarSharma"

//Read Operation.

RCString str3("RajeevKumarNayan");

RCString str4(str3);

cout << str3[0] << endl; //Read, Unfortunately in operator[] call, we cannot

//determine whether it is called by read or write

//operation.

str3.showStr(); //str3->refCount=1

str4.showStr(); //str4->refCount=1

//Still having seperate copy of string //"RajeevKumarNayan" which is illogical.

const RCString str5("ThisIsConstString");

str5.showStr();

char ch1 = str5[3]; //Non Const RCString Read.

str5.showStr(); //No address change after read, because of const

RCString str6("ThisIsNonConstString")

str6.showStr();

char ch2 = str6[1]; //String changed while reading.

str6.showStr(); //Address changed.

getchar();

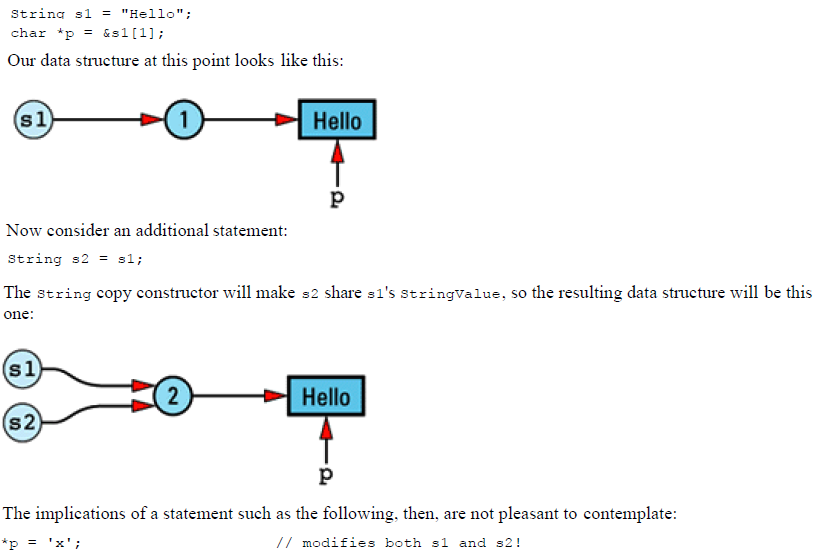
return 0;

}

**Pointers, References, and Copy-on-Write:**

This idea — that of sharing a value with other objects until we must write on our own copy of the value — has a long and distinguished history in Computer Science, especially in operating systems, where processes are routinely allowed to share pages until they want to modify data on their own copy of a page. The technique is common enough to have a name: ***copy-on-write***.

Copy-on-write allows us to preserve both efficiency and correctness.



There is no way the String copy constructor can detect this problem, because it has no way to know that a pointer into s1's ***StringValue*** object exists. And this problem isn't limited to pointers, it would exist in previous example, when we call to ***RCString*** non-const operator[].

So how we resolve this problem?

Its implement is not difficult, but it can reduce the amount of value sharing between objects. We can add a flag to each StringValue object indicating whether that object is shareable. Turn the flag on initially (the object is shareable), but turn it off whenever the

non-const operator[] is invoked on that object. Once the flag is set to false, it stays for forever.

class String {

private:

struct StringValue {

int refCount;

bool shareable; // added this

char \*data;

StringValue(const char \*initValue);

~StringValue();

};

...

};

String::StringValue::StringValue(const char \*initValue): refCount(1),

shareable(true) // added this

{

data = new char[strlen(initValue) + 1];

strcpy(data, initValue);

}

There is not much changes required to implements this. Of course, String's member functions must be updated to take the shareable field into account. Here's how the copy constructor implemented:

String::String(const String& rhs)

{

if (rhs.value->shareable) {

value = rhs.value; //If sharable is true then only

++value->refCount; //increase the ref cnt.

}

else {

value = new StringValue(rhs.value->data);

//Otherwise creates new copy.

}

}

And Finally, The non-const version of operator[] would be the only function to set the shareable flag to false.

char& String::operator[](int index)

{

if (value->refCount > 1) {

--value->refCount;

value = new StringValue(value->data);

}

value->shareable = false; // add this

return value->data[index];

}

Note: If we use the proxy class technique of to distinguish read usage from write usage in operator[], then we can reduce the number of StringValue objects that must be marked un-shareable.

**A Reference-Counting Base Class:**

Reference counting is not only useful for String object, but for any class in which different objects may have common values. So, if we could somehow write the reference counting code in a context-independent manner.

The first step is to create a base class, RCObject, for reference-counted objects. Any class wishing to take advantage of automatic reference counting must inherit from this class.

RCObject encapsulates the reference count itself, as well as functions for incrementing and decrementing that count. It also contains the code for destroying a value when it is no longer in use, i.e., when its reference count becomes 0. Finally, it contains a field that keeps track of whether this value is shareable, and it provides functions to query this value and set it to false. There is no need for a function to set the shareability field to true, because all values are shareable by default. As noted above, once an object has been tagged un-shareable, there is no way to make it shareable again. RCObject's class definition looks like this:

class RCObject {

public:

RCObject();

RCObject(const RCObject& rhs);

RCObject& operator=(const RCObject& rhs);

virtual ~RCObject() = 0;

void addReference();

void removeReference();

void markUnshareable();

bool isShareable() const;

bool isShared() const;

private:

int refCount;

bool shareable;

};

RCObjects can be created and destroyed; they can have new references added to them and can have current references removed; their shareability status can be queried and can be disabled; and they can report whether they are currently being shared.

RCObject::RCObject(): refCount(0), shareable(true) {}

RCObject::RCObject(const RCObject&): refCount(0), shareable(true) {}

RCObject& RCObject::operator=(const RCObject&){

return \*this;

}

RCObject::~RCObject() {} // virtual dtors must always be implemented,

// even if they are pure virtual and do nothing

void RCObject::addReference() { ++refCount; }

void RCObject::removeReference(){

if (--refCount == 0)

delete this;

}

void RCObject::markUnshareable(){

shareable = false;

}

bool RCObject::isShareable() const{

return shareable;

}

bool RCObject::isShared() const{

return refCount > 1;

}

we set refCount to 0 inside both constructors. Will see why?

To take advantage of our new reference-counting base class, we modify StringValue to inherit its reference counting capabilities from RCObject.