

Taller de Programación I (75.42)

Clase 5: Templates y STL

Leandro Lucarella
llucax+7542@gmail.com
Departamento de Computación, Facultad de Ingeniería
Universidad de Buenos Aires

Revisión: 1

1. Introducción

Las *templates* proveen a C++ la posibilidad de realizar [Programación Genérica](#). La programación genérica permite programar algoritmos paramétricos, utilizando (principalmente) tipos como parámetros, de manera tal de que un algoritmo pueda ser utilizado intacto para distintos tipos de datos, que comparten una interfaz común.

A simple vista pareciera que esto no ofrece ninguna ventaja sobre la programación orientada a objetos, con la cual se puede lograr el mismo efecto a través de una jerarquía de clases (o interfaces), razón por la cual también suele llamarse a las *templates* **polimorfismo estático**.

Además las *templates* hacen prácticamente obsoletas las funciones *macro* del precompilador, ya que al ser resueltas en tiempo de compilación son tan eficientes como estas, pero al ser parte del lenguaje C++, pueden hacer uso de su estado interno y son, entre otras cosas, *type-safe*.

STL significa *Standard Template Library* (Biblioteca estándar de plantillas, o *templates*). A la biblioteca estándar de C, C++ agrega una biblioteca estándar de tipos y algoritmos parametrizados, entre ellos búsqueda binaria (y otros tipos de búsquedas, ordenamiento, máximo, mínimo, reemplazo, contenedores, *strings*, números complejos, *streams*, etc.

2. Funciones parametrizadas

C++ permite parametrizar funciones de varias maneras, veamos algún ejemplo:

```
template < typename Tipo >
Tipo min(const Tipo& x, const Tipo& y)
{
    return x > y ? y : x;
}
```

Consejo

Es muy recomendable dejar espacios entre los “corchetes angulares” (< y >) porque en el caso de anidar *templates*, si no se hace esto, el compilador los va a confundir con el operador de *shift*: >>, por ejemplo en este caso `std::map<int, std::list<double>`, al final nos queda el operador >> y el compilador a veces da un mensaje de error muy extraño.

La palabra clave `template` permite declarar (y/o definir) una función parametrizada. A continuación se indican los parámetros de la *template* (no confundir con los parámetros de la función). Para indicar

que toma como parámetro un **tipo** se pueden utilizar las palabras reservadas `typename` o `class` (indistintamente, tienen el mismo significado **en este contexto**), es decir, podría haber escrito `template <class T>` y el resultado sería el mismo. Recomendamos utilizar la palabra `typename` para enfatizar que el parámetro no tiene que ser necesariamente una clase, puede ser cualquier otro tipo nativo del compilador (como `int`, `char` o incluso un `enum`). Finalmente `Tipo`, es el nombre de nuestro parámetro.

A continuación sigue la declaración y definición de una función común y corriente, con la salvedad de que utilizaremos `Tipo` como si fuera el nombre de un tipo existente (recordar que es un parámetro, podría no existir ningún tipo llamado `Tipo`).

Ahora es posible utilizar la función `min()` con cualquier tipo, por ejemplo `int` o `double` o una clase propia.

Nota

Realmente se podrá utilizar cualquier tipo? O nuestra función impondrá alguna restricción?

3. Instanciación

Las templates en realidad no existen (*there is no spoon*), no al menos hasta que son **instanciadas**. Esto significa que una *template*, sin tener el tipo definido, no puede generar código, porque es imposible para el compilador generar código que utiliza tipos de los cuales todavía no conoce el tamaño (en el caso de la función `min()`, cuantos bytes debería apilar en el *stack* al llamarse la función para pasar los parámetros?). Es por esto que si una *template* no se usa, **no genera código**, no existe.

De esto se desprenden varios fenómenos particulares, a saber:

- Debe haber un mecanismo para instanciar *templates*.
- Las templates deben estar **definidas** (no solo declaradas) en todos los archivos que la usen (veremos esto en más detalle más adelante).

Para **instanciar** una *template* hay dos formas. La más común es simplemente utilizarla, por ejemplo:

```
int i = 5;
int j = min(i, 7);
```

Instancia la template `int min(const int&, const int&)`, es decir, instancia nuestra template con `Tipo = int`. Esto **genera código**, el código para la función `int min(const int&, const int&)`. Si luego escribimos:

```
double d = min(1.0, 4.2);
```

Nuevamente se instancia la template, esta vez con `Tipo = double`, generando **otra función**: `double min(const double&, const double&)` (recordar que en C++ se permite la sobrecarga de funciones, funciones con el mismo nombre pero distinta cantidad y/o tipo de parámetros, por lo que una función en realidad es identificada por su nombre más sus parámetros). Es decir, nuestro código objeto ahora tiene **dos** funciones `min()`. Y antes de empezar esta sección (antes de utilizarla), no tenía **ninguna**. Qué pasa si ahora escribimos en nuestro programa:

```
int k = min(i, j);
```

No se crea más código del que ya había, se utiliza la *template* ya instanciada con `Tipo = int`. También es posible instanciar una *template* explícitamente:

```
int l = min< float >(i, j);
```

Si bien `min()` toma como parámetros dos `int`, se instancia la *template* `float min(const float&, const float&)`, ya que se especificó explícitamente que `Tipo = float`. Al igual que si hubiera utilizado una función no parametrizada, el compilador se encargará de *castear* los `int` a `float` antes de pasarle los parámetros a la función.

¡Atención!

Es muy importante que se comprenda este concepto antes de continuar con la lectura.

4. Organización en archivos

Como se comentó anteriormente, como el código de una *template* no es generado hasta que esta no se instancia, y como, generalmente, la instanciación viene asociada al uso, es necesario asegurarse que la **definición** de una *template* esté disponible para todos los archivos en donde ésta se use. Por lo tanto la práctica más usual es tanto **declarar** como **definir** las templates en los `.h`. En nuestro ejemplo, esto puede hacerse directamente poniendo el primer fragmento de código en un `.h`, o bien separándolo **declaración** y **definición**, pero siempre dejando todo en el `.h`:

```
template < typename Tipo >
Tipo min(const Tipo& x, const Tipo& y);

template < typename Tipo >
Tipo min(const Tipo& x, const Tipo& y)
{
    return x > y ? y : x;
}
```

Para una sola función (y tan simple), esto parece absurdo, pero cuando se declaran varias *templates* en un `.h`, y más aún cuando su implementación no es trivial, puede ser mucho más claro declarar todas las funciones primero, y definir las luego (incluso separando definición y declaración en archivos `.h` diferentes, asegurándose que al incluir el archivo con las declaraciones, también se incluya el de definiciones). Esto cobra aún más sentido cuando se parametrizan clases, ya que si no se separan declaración de definición, todos los métodos de la clase pasan a ser `inline` implícitamente.

5. Clases parametrizadas

La parametrización de clases es muy útil, en especial, al desarrollar *contenedores*; es decir, clases que sirven para almacenar otros objetos (como listas, vectores, tablas de hash, etc.).

Veamos un ejemplo clásico (completo) de como encapsular un array estático parametrizándolo:

```
#include <iostream>
#include <algorithm>
#include <cassert>
#include <cstdlib>

template <
    typename Tipo,          // Tipo que almacena el array
    size_t Tamanio = 5 > // Parámetro numérico y por omisión
struct Array
{
    // Tipos y constantes: para facilitar la programación genérica
    typedef Tipo value_type;
    enum { MAX_SIZE = Tamanio };
    // Constructores
    Array() {} // default, inline
    Array(const Tipo* array, size_t tamanio);
    // Acceso a un elemento a través de un índice para escritura
    Tipo& operator[] (size_t indice);
};
```

```

        // Acceso a un elemento a través de un índice para lectura
        const Tipo& operator[] (size_t indice) const;
        // Acceso a un elemento a través con chequeo de "bordes" (bound-check)
        Tipo& at(size_t indice);
        const Tipo& at(size_t indice) const;
        // Obtención de tamaño máximo
        size_t capacity() const { return MAX_SIZE; } // Este método es "inline"
        // Conversión a un array de otro tipo
        template < typename OtroTipo >
        Array< OtroTipo, Tamanio > as() const;
        // Array estático
        Tipo _array[Tamanio];
};

template < typename Tipo, size_t Tamanio >
Array< Tipo, Tamanio >::Array(const Tipo* array, size_t tamanio)
{
    for (size_t i = 0; i < tamanio; ++i) _array[i] = array[i];
}

template < typename Tipo, size_t Tamanio >
Tipo& Array< Tipo, Tamanio >::operator[] (size_t indice)
{
    return _array[indice];
}

template < typename Tipo, size_t Tamanio >
const Tipo& Array< Tipo, Tamanio >::operator[] (size_t indice) const
{
    return _array[indice];
}

template < typename Tipo, size_t Tamanio >
Tipo& Array< Tipo, Tamanio >::at(size_t indice)
{
    assert(indice < MAX_SIZE);
    return _array[indice];
}

template < typename Tipo, size_t Tamanio >
const Tipo& Array< Tipo, Tamanio >::at(size_t indice) const
{
    assert(indice < MAX_SIZE);
    return _array[indice];
}

// Para esto, Tipo debe ser implícitamente convertible a OtroTipo
template < typename Tipo, size_t Tamanio >
template < typename OtroTipo >
Array< OtroTipo, Tamanio > Array< Tipo, Tamanio >::as() const
{
    Array< OtroTipo, Tamanio > otro;
    std::copy(_array, _array + Tamanio, otro._array); // algoritmo de STL
}

```

```

        return otro;
    }

    int main(int argc, char* argv[])
    {
        typedef Array< int > ArrayInt; // Para escribirlo más fácil (5 elem.)
        typedef Array< char*, 10 > ArrayCharp; // con 10 elementos
        assert(argc > ArrayInt::MAX_SIZE);
        ArrayInt arrint;
        ArrayCharp args(argv+1, argc-1);
        for (size_t i = 0; i < argc-1; ++i)
        {
            std::cout << i << ": " << args.at(i); // con bound-check
            // arrint lo accedemos con chequeo de bordes, args no es
            // necesario porque ya chequeamos bordes en la línea anterior.
            arrint.at(i) = atoi(args[i]);
            std::cout << " -> " << arrint[i] << "\n"; // sin bound-check
        }
        // convertimos a array de double
        Array< double > arrdbl = arrint.as< double >();
        std::cout << "Con doubles:\n";
        std::cout.precision(2);
        for (size_t i = 0; i < arrdbl.capacity(); ++i)
        {
            std::cout << i << ": " << std::fixed << arrdbl[i] << "\n";
        }
        return 0;
    }
}

```

Analicemos el ejemplo por partes. Primero observamos un tipo de parámetro para una *template* nuevo: `size_t Tamano = 5`. Hay dos conceptos nuevos en este parámetro: el tipo (`size_t`) y el valor por omisión (= 5). Una *template* puede recibir también como parámetro cualquier tipo nativo del compilador, y su valor puede ser utilizado como una constante resuelta en tiempo de compilación dentro de la *template* (esto permite utilizarla como tamaño para un array estático, como vemos en el ejemplo).

Luego podemos ver un `typedef` y un `enum` con un comentario que indica que facilitan la programación genérica, explica esto excede el objetivo de este documento pero pueden ampliar buscando en Internet o en la bibliografía de la materia. Esta técnica es ampliamente utilizada por la STL.

A continuación se declaran un par de constructores, siendo el constructor *default inline* (es decir, el compilador hará lo posible por no generar una llamada a función cuando se utilice, en este caso como es vacío, seguramente podrá hacerlo). El método `capacity()` también es *inline*.

Luego se declaran dos variantes del `operator[]`, una para ser utilizado de forma constante (para lectura solamente) y otro para modificación, acompañado cada uno por una versión (`at()`) que incluye un rudimentario chequeo de bordes (*bound-checking*) implementado via la *macro* `assert()` de C, para asegurarse de no *pisar* memoria si nos vamos de los límites del array. Esto también se incluye a modo de ejemplo por ser una técnica utilizada por los contenedores de la STL (sólo que éstos utilizan excepciones para reportar errores).

Finalmente nos encontramos con un método parametrizado, dentro de una clase parametrizada (`as()`), que sirve para realizar una conversión genérica de un tipo de array a otro. Ambos arrays tienen el mismo tamaño, pero contienen tipos de datos arbitrariamente diferentes (mientras la conversión implícita sea posible).

Luego nos encontramos con el único atributo de esta clase, que utiliza ambos parámetros, el `Tipo` y el `Tamano` (no se utilizó control de acceso, `private`, por motivos didácticos).

Después de la declaración de la clase (**struct** en este caso, por simplicidad), están todas las definiciones de los métodos no *inline*. Se puede prestar particular atención a la definición del método **as()** que tiene 2 particularidades interesantes:

1. La forma de definir un método parametrizado de una clase parametrizada (como se puede observar, esto requiere de dos palabras clave **template**).
2. La utilización de un algoritmo de la STL para copiar los datos de un array a otro (como adelanto de los algoritmos de la STL).

El resto del programa es una función **main()** que demuestra las capacidades de nuestra clase genérica. Primero muestra lo conveniente que es la utilización de **typedef** al utilizar *templates*, además de demostrar la instanciación de una template con un parámetro por omisión. Luego muestra la utilidad de haber definido el **enum** con **MAX_SIZE**, porque de otra forma, para escribir el **assert()**, hubiera sido necesario ver cual es el valor del parámetro por default de la template y ponerlo a mano. Esto además de ser engorroso, hace el código inmantenible, porque si un día el autor de la *template* cambia ese valor por default, introduce un *bug* en nuestro código (o hay que estar pendiente todo el tiempo de los cambios). Después se instancia un objeto de la clase **Array<char*, 10 >**, inicializándolo con el array **argv** utilizando el constructor.

Importante

Notar que el nombre de la clase es **Array<char*, 10 >**, no **Array** solo, **Array** es una *template*, no una clase.

Luego se copian los valores de un array a otro (convirtiendo los valores de texto a **int**), haciendo *bound-checking*, y finalmente se convierte (utilizando el método **as()**) el array de **int** a uno de **double**, imprimiendo el resultado. También se puede observar en estos pasos algunas funciones de formateo de los *streams* de la STL.

6. Especialización

Una característica muy conveniente de las *templates* de C++, es la especialización. Esto permite definir un comportamiento específico para un tipo de datos particular de una *template*.

Son varias las veces donde es posible proveer una implementación más eficiente (o simplemente más *correcta*) para un dato en particular de lo que puede proveer un algoritmo genérico. Un ejemplo de esto puede ser un array de **bool** (que puede ser implementado utilizando bits). Incluso nuestra clase **Array** se puede beneficiar mucho de la especialización. Que pasa si hiciera una copia del array **args**, que almacena punteros a **char**? Que pasa si se liberaran esos punteros? La copia quedaría apuntando a memoria liberada, un error común.

Para evitar esto, se podría hacer una versión del **Array** especializada, para que en el constructor aloque memoria y en el destructor la libere. En realidad con nuestro ejemplo esto no puede hacerse por el parámetro por omisión, pero si no tuvieramos el segundo parámetro (**Tamano**), podríamos especificar la *template* para que tenga una implementación particular para **char***, de esta manera:

```
template <>
struct Array< char* >
{
    // Tipos y constantes: para facilitar la programación genérica
    typedef char* value_type;
    ////////// más miembros pero especificados con char* //////////
    // Array estático
    char* _array[Tamano];
};
```

Una desventaja, es que hay que redeclarar la clase completa (incluso ambas clases podrían no tener los mismos métodos; recordar que cada **instancia** de una *template* es una clase **distinta**). Pero hay otra desventaja aun mayor en este caso. Qué pasa si quisieramos ahora utilizar un `Array<int* >`? Habría que hacer toda la especialización de nuevo para `int*`? Eso sería muy engorroso e iría en contra de otro de los objetivos de las *templates*: reutilización de código. Para esto es que existe la **especialización parcial**. La idea detrás de este concepto es poder especializar una clase para punteros, pero no un puntero en particular, sino cualquier puntero. Para hacer esto podríamos declarar:

```
template < typename Tipo >
struct Array< Tipo* >
{
    // Tipos y constantes: para facilitar la programación genérica
    typedef Tipo* value_type;
    ////////// más miembros pero especificados con Tipo* //////////
    // Array estático
    Tipo* _array[Tamano];
};
```

De esta manera, al crear un objeto del tipo `Array<int* >` se instanciaría una sola *template*, que se compartiría con otras instancias que utilicen punteros, como `Array<double* >`.

No vamos a entrar más en detalle con respecto a especialización porque es un tema complejo que excede el objetivo de esta clase. Siempre pueden ampliar en internet o en la bibliografía recomendada por la cátedra.

7. STL

STL significa Standard Template Library (en castellano algo así como Biblioteca Estándar de Plantillas) y es parte de la biblioteca estándar de C++ (en conjunto con las funciones de la biblioteca estándar de C). El objetivo principal de dicha biblioteca es albergar algoritmos y tipos de datos (en particular lo que se conoce como contenedores) tan básicos que prácticamente todo programa no trivial haga uso de al menos alguno de ellos.

A diferencia de otras bibliotecas estándar (como la de Java, Python o .NET), no pretende tener respuesta a todas las necesidades que alguna vez puede tener un programador en su vida, sólo busca crear una base, un mínimo común múltiplo, que sirva para construir bibliotecas o programas más complejos y específicos. En este sentido es muy similar a la biblioteca estándar de C, sólo que la STL va un paso más allá, implementando estructuras de datos básicas (listas, árboles, heaps, vectores dinámicos, etc.). Por lo tanto uno de los grandes objetivos de diseño de la STL es la genericidad. La STL siempre evitará tomar decisiones por el usuario (o al menos siempre habrá forma de cambiar el comportamiento por omisión).

Otra característica es la eficiencia. C++ pretende ser un lenguaje apto para escribir programas con grandes requerimientos de performance (sistemas de tiempo real por ejemplo). Es por esto que la STL sacrifica facilidad de uso por performance. Si bien no es extremadamente compleja de usar, hay construcciones que pareciera que podrían ser más simples y muchas veces no lo son para poder dar lugar a optimizaciones (un ejemplo clásico es que los métodos tipo `pop()` no retornan el elemento eliminado, hay que usar el método `top()` para obtener el tope y luego hacer el `pop()` para *sacarlo*). Otra particularidad en este sentido es que están especificado hasta los ordenes de complejidad de los algoritmos de la STL, permitiendo tener un control bastante fino sobre la performance incluso sin conocer la implementación.