C++ Tricks 3.2 标号、goto，以及 switch 的实现 ☒

从 farseerfc.wordpress.com 导入

3.2 标号、goto，以及 switch 的实现

goto 语句及标号 (label) 是最古老的 C 语言特性，也是最早被人们抛弃的语言特性之一。像汇编语言中的 jmp 指令一样，goto 语句可以跳转到同一函数体中任何标号位置：

```
void f()
{int i=0;
Loop: //A label
```

```
++i;
if(i<10)goto Loop; //Jump to the label
}
```

在原始而和谐的早期 Fortran 和 Basic 时代，我们没有 if then else，没有 for 和 while，甚至没有函数的概念，一切控制结构都靠 goto(带条件的或无条件的) 构件。软件工程师将这样的代码称作“意大利面条”代码。实践证明这样的代码极容易造成混乱。

自从证明了结构化的程序可以做意大利面条做到的任何事情，人们就开始不遗余力地推广结构化设计思想，将 goto 像猛兽一般囚禁在牢笼，标号也因此消失。

标号唯一散发余热的地方，是在 switch 中控制分支流程。

很多人不甚了解 switch 存在的意义，认为它只是大型嵌套 if then else 结构的缩略形式，并且比 if 语句多了很多“不合理”的限制。如果你了解到 switch 在编译器内部的实现机制，就不难理解强加在 switch 之上的诸多限制，比如 case 后只能跟一个编译期整型常量，比如用 break 结束每一个 case。首先看一个 switch 实例：

```
switch (shape.getAngle())
{
case 3: cout<<”Triangle”;break;
case 4: cout<<”Square”;break;
case 0:case1: cout<<”Not a sharp!”;break;
default: cout<<”Polygon”;
}
```

任何程序员都可以写出与之对应的 if 结构：

```
int i= getAngle(shape);
```

if (i==3) cout<<”Triangle”;

else if(i==4) cout<<”Square”;

else if(i==0||i==1) cout<<”Not a sharp!”;

else cout<<”Polygon”;

看起来这两段代码在语义上是完全一样的，不是么？

不！或许代码的执行结果完全一样，但是就执行效率而言，switch 版本的更快！

要了解为什么 switch 的更快，我们需要知道编译器是怎样生成 switch 的实现代码的：

首先，保留 switch 之后由{}括起来的语具体，仅将其中 case、default 和 break 替换为真正的标号：

switch (getAngle(shape))

{

_case_3: cout<<”Triangle”;goto _break;

_case_4: cout<<”Square”; goto _break;

_case_0:_case_1: cout<<”Not a sharp!”; goto _break;

_default: cout<<”Polygon”;

_break:

}

随后，对于所有出现在 case 之后的常量，列出一张只有 goto 的跳转表，其顺序按 case 后的常量排列：

goto _case_0;

goto _case_1;

goto _case_3;

goto _case_4;

然后，计算 case 之后的常量与跳转表地址之间的关系，如有需要，在跳转表中插入空缺的项目：

100105: goto _case_0;

100110: goto _case_1;

100115: goto _default; //因为没有 case 2,所以插入此项以条转到 default

100120: goto _case_3;

100125: goto _case_4;

假设一个 goto 语句占用 5 个字节，那么在本例中，goto 的地址 =case 后的常量*5+100105

之后，生成跳转代码，在其余条件下跳转至 default，在已知范围内按照公式跳转，全部的实现如下：

{

int i= getAngle(shape);

if (i<0||i>=5)goto _default;

i=i*5+100105; //按照得出的公式算出跳转地址

goto i; //伪代码，C 中不允许跳转到整数，但是汇编允许

100105: goto _case_0;

100110: goto _case_1;

100115: goto _default;

100120: goto _case_3;

```
100125: goto _case_4;

_case_3: cout<<”Triangle”;goto _break;

_case_4: cout<<”Square”; goto _break;

_case_0:_case_1: cout<<”Not a sharp!”; goto _break;

_default: cout<<”Polygon”;

_break:

}
```

经过这样处理整个 switch 结构，使得无论 switch 后的变量为何值，都可以通过最多两次跳转到达目标代码。相比之下 if 版本的代码则采用线性的比较和跳转，在 case 语句很多的情况下效率极低。

由此, 我们也可以知道, 为什么 case 后跟的一定是编译期整型常数，因为编译器需要根据这个值制作跳转表。我们可以明白为什么 case 与 case 之间应该用 break 分隔，因为编译器不改变 switch 语句体的结构，case 其本身只是一个具有语义的标号而已，要想跳出 switch，就必须用 break 语句。