Procesadores de Lenguaje

Tema 2.2. Ampliación del lenguaje y su procesador: Optimización de la traducción



Profesor

Federico Peinado

Elaboración del material

José Luis Sierra Federico Peinado

Ingeniería en Informática
Facultad de Informática — Universidad Complutense de Madrid
Curso 2009-2010

Optimización de la traducción

- ☐ En general, **optimizar un traductor** es hacer que genere programas *equivalentes pero más eficientes* en tiempo de ejecución o consumo de memoria
 - Disciplina fundamental en el diseño de compiladores reales
- Hay distintos aspectos que pueden ser optimizados:
 - Las estrategias de traducción a seguir
 - ☐ El código objeto una vez generado
 - ☐ El propio código fuente
- Nosotros sólo veremos algunos casos básicos, aunque Aho et al. enseñan métodos más complejos y sistemáticos



Traducción de instrucciones if-then-else sin bloque else

□ Cuando una instrucción if-then-else no tiene bloque else se genera siempre un código en el bloque then con un salto totalmente innecesario a la instrucción siguiente

Ejemplo:

x = Mem[0]

. . .

```
if x = 5 then x := x + 1;
```

```
apila-dir(0)
apila(5)
igual
ir-f(9)
apila-dir(0)
apila(1)
suma
desapila-dir(0)
ir-a(9)

9
```



■ La traducción se puede optimizar generando o no el código de ese salto según el bloque else esté vacío o no

Ampliación de la propuesta: if-then-else sin bloque else

☐ En nuestras ecuaciones semánticas de la gramática de atributos esta optimización queda así:

```
IIf ::= if Exp then I PElse
  IIf.cod = Exp.cod | ir-f(PElse.etgi) | I.cod | |
            PElse.cod
 Exp.etgh = IIf.etgh
  I.etgh = Exp.etg + 1
 PElse.etqh = I.etq
  IIf.etq = PElse.etq
PElse ::= else T
 PElse.cod = ir-a(I.etq) | I.cod
  I.etqh = PElse.etqi = PElse.etqh + 1
  PElse.etq = I.etq
PElse ::= \lambda
 PElse.cod = \lambda
  PElse.etq = PElse.etqi = PElse.etqh
```



Ampliación de la propuesta: if-then-else sin bloque else

☐ Y en los esquemas de traducción optimizados queda así:

```
global cod, etq;
...
IIf::= {var etqaux}
    if Exp then
        {emite(ir-f(?));
        etqaux ← etq;
        etq ← etq + 1}
    I
    PElse(out etqi)
    {parchea(etqaux, etqi)}
```

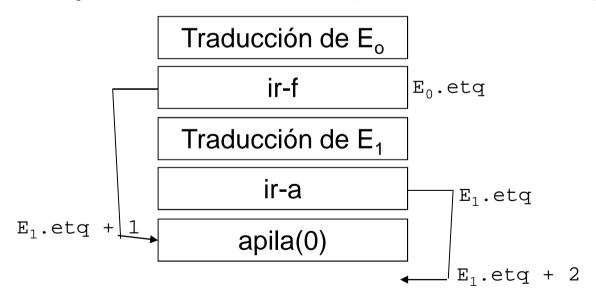


Ampliación de la propuesta: if-then-else sin bloque else

```
global \ cod, \ etq;
...
PElse(out \ etqi) ::= \{var \ etqaux\}
else
\{emite(ir-a(?));
etqaux \leftarrow etq;
etq \leftarrow etq + 1;
etqi \leftarrow etq\}
I
\{parchea(etqaux, \ etq)\}
PElse(out \ etqi) ::= \lambda
\{etqi \leftarrow etq\}
```

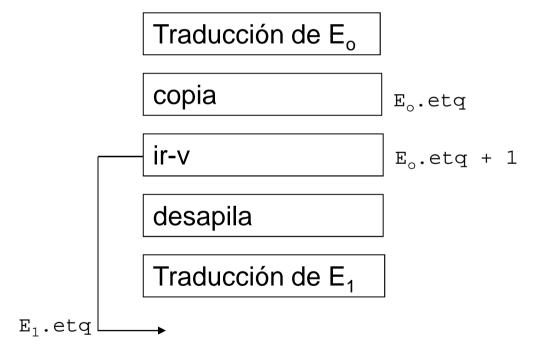


- □ Es habitual traducir las expresiones booleanas para que se calcule solamente el valor de aquellos operandos imprescindibles para conocer el resultado de la expresión
 - Usamos las mismas ideas de las sentencias condicionales
- Esquema de E₀ and E₁
 - Si E₀ es falso, todo vale falso; y si no, vale lo que valga E₁





- Esquema de E₀ or E₁
 - Si E₀ no es falso, todo vale lo que valga E₀ (verdadero); y si no, vale lo que valga E₁





- * copia duplica la cima de la pila (lo copia otra vez apilándolo encima)
- * ir-v realiza el salto si hay verdad en la cima de la pila (lo contrario de ir-f)
- * desapila descarta la cima de la pila

☐ copia

```
Pila [CPila + 1] ← Pila[CPila]
CPila ← CPila + 1
CProg ← CProg + 1
```

 \square ir-v(i)

```
si Pila[CPila] \neq 0

CProg \leftarrow i

si no

CProg \leftarrow CProg + 1

fsi

CPila \leftarrow CPila - 1
```

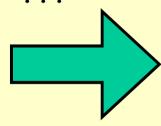
☐ desapila



□ Ejemplo:

x = Mem[0]
y = Mem[1]
z = Mem[2]
u = Mem[3]
v = Mem[4]
w = Mem[5]
r = Mem[6]

(x=y) and (y=z) or (u=v) and (w=r)





apila-dir(0) apila-dir(1) igual ir-f(8) 3 apila-dir(1) apila-dir(2) 5 igual 6 ir-a(9) apila(0) copia 9 ir-v(21)10 desapila 11 apila-dir(3) 12 13 apila-dir(4) igual 14 ir-f(20) 15 apila-dir(5) 16 17 apila-dir(6) igual 18 ir-a(21) 19 apila(0) 20 21

Ampliación de la propuesta: Expresiones booleanas en circuito corto

- □ Para facilitar esta generación de código optimizado es conveniente modificar la gramática incontextual para discriminar explícitamente los operadores lógicos del resto
 - ☐ Añadiendo un par de producciones para los operadores:

```
ExpS ::= ExpS or Term
...
Term ::= Term and Fact
```

☐ Eliminando los operadores de las otras producciones:

```
OpAd ::= + | - | or
...
OpMul ::= * | / | div | mod | and
```



Ampliación de la propuesta: Expresiones booleanas en circuito corto

□ De esa forma podemos definir las partes correspondientes a la traducción optimizada en la gramática de atributos:

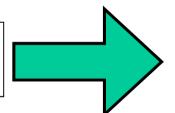
```
ExpS ::= ExpS or Term
  ExpS_0.cod = ExpS_1.cod \mid copia \mid ir-v(Term.etq) \mid 
                desapila | Term.cod
  ExpS_1.etqh = ExpS_0.etqh
  Term.etqh = Exps_1.etq + 3
  ExpS<sub>o</sub>.etq = Term.etq
Term ::= Term and Fact
  Term<sub>o</sub>.cod = Term<sub>1</sub>.cod || ir-f(Fact.etq + 1) ||
                Fact.cod | ir-a(Fact.etq + 2) | apila(0)
  Term<sub>1</sub>.etqh = Term<sub>0</sub>.etqh
  Fact.etqh = Term_1.etq + 1
  Term_o.etq = Fact.etq + 2
```



☐ La optimización propuesta para la traducción de las expresiones booleanas produce saltos consecutivos ineficientes al anidar operadores lógicos

Ejemplo:

$$(x=5)$$
 or $(x > 6)$ or $(x < -1)$

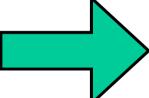




apila-dir(0) apila(5) iqual copia ir-v(9)desapila apila-dir(0) apila(6) mayor copia ir-v(15)10 11 12 apila(-1) 13 14 menor 15

 $x \equiv Mem[0]$

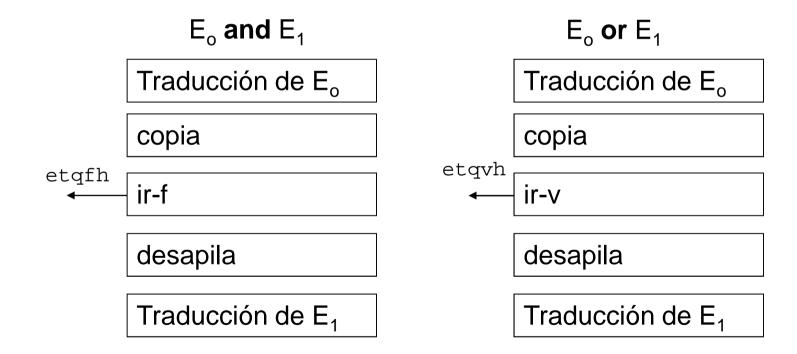
- ☐ La traducción se puede optimizar, por segunda vez, haciendo que los saltos lleguen al final del anidamiento (que sean "saltos completos") $x \equiv Mem[0]$
- Ejemplo:





apila-dir(0) apila(5) iqual copia ir-v(15)desapila mayor 10 11 12 13 14 menor 15

- ☐ La idea ahora es fijar *siempre* las etiquetas de los saltos condicionales **desde las estructuras de nivel superior**
 - Los saltos **ir-f** / **ir-v** serán *completos*: tendrán como dirección de destino el valor de un atributo heredado etqfh / etqvh que indica el final del anidamiento de operadores **and** / **or**



Así se expresan los valores de etqfh y etqvh en las ecuaciones semánticas de las expresiones y de cualquier otra estructura que utilice expresiones:

```
IWhile::= while Exp do I
   Exp.etqfh = Exp.etqvh = Exp.etq

Iif:: if Exp then I PElse
   Exp.etqfh = Exp.etqvh = Exp.etq

Exp.etqfh = Exp.etqvh = Exp.etq

ExpS_0.etqfh = ExpS_0.etqvh = ExpS_0.etq
   ExpS_1.etqfh = ExpS_1.etqvh = ExpS_1.etq

***
```



```
Exp ::= ExpS
  ExpS.etgfh = Exp.etgfh
  ExpS.etqvh = Exp.etqvh
ExpS ::= ExpS OpAd Term
  ExpS_1.etqfh = ExpS_1.etqvh = ExpS_1.etq
  Term.etgfh = Term.etgvh = Term.etg
ExpS ::= Term
  Term.etgfh = ExpS.etgfh
  Term.etqvh = ExpS.etqvh
Term ::= Term OpMul Fact
  Term<sub>1</sub>.etqfh = Term<sub>1</sub>.etqvh = Term<sub>1</sub>.etq
  Fact.etqfh = Fact.etqvh = Fact.etq
Term ::= Fact
  Fact.etqfh = Term.etqfh
  Fact.etqvh = Term.etqvh
Fact ::= not Fact
  Fact<sub>1</sub>.etqfh = Fact<sub>1</sub>.etqvh = Fact<sub>1</sub>.etq
Fact ::= (Exp)
  Exp.etqfh = Fact.etqfh
  Exp.etqvh = Fact.etqvh
```



```
ExpS ::= ExpS or Term
  ExpS_0.cod = ExpS_1.cod \mid copia \mid
                ir-v(ExpS<sub>0</sub>.etqvh) || desapila ||
                Term.cod
  ExpS_1.etqfh = ExpS_1.etq + 2
  Term.etqfh = Term.etq
  ExpS_1.etqvh = ExpS_0.etqvh
  Term.etqvh = ExpS_0.etqvh
Term ::= Term and Fact
  Term<sub>0</sub>.cod = Term<sub>1</sub>.cod | copia |
                ir-f(Termo.etqfh) | desapila | |
                Fact.cod
  Term_1.etqfh = Term_0.etqfh
  Fact.etqfh = Term_0.etqfh
  Term_1.etqvh = Term_1.etq + 2
  Fact.etqvh = Fact.etq
```



- ☐ La anterior gramática de atributos **no es l-atribuida**: ¡hay muchas categorías sintácticas que *heredan de sí mismas*!
 - □ Por ejemplo: Exp.etqfh = Exp.etqvh = Exp.etq
- □ Esto es un problema que requiere volver a utilizar la idea del parcheo (ahora en la propia gramática de atributos), parcheando muchos saltos, todos con el mismo destino
- □ Para ello añadimos dos atributos sintetizados que almacenan listas de instrucciones pendientes de parchear
 - Los saltos ir-f con valor indefinido se guardarán en listaf
 - Los saltos ir-v con valor indefinido se guardarán en listav
- ☐ Y también añadimos una función semántica parchea que realiza el parcheo "masivo" de las dos listas de instrucciones, dada la dirección de destino que hay que usar en las instrucciones de salto de cada una de ellas
 - parchea (cod, listaf, listav, etqf, etqv)



☐ Así quedan las ecuaciones semánticas de la traducción:



```
Exp ::= ExpS OpComp ExpS
  Exp.cod = parchea(ExpS_0.cod, ExpS_0.listaf, ExpS_0.listav,
                       ExpS_0.etq, ExpS_0.etq)
             parchea(ExpS<sub>1</sub>.cod, ExpS<sub>1</sub>.listaf, ExpS<sub>1</sub>.listav,
                       ExpS_1.etq, ExpS_1.etq) | OpComp.op
  Exp.listaf = Exp.listav = []
Exp ::= ExpS
  Exp.listaf = ExpS.listaf
  Exp.listav = ExpS.listav
ExpS ::= ExpS OpAd Term
  ExpS_0.cod = parchea(ExpS_1.cod, ExpS_1.listaf, ExpS_1.listav,
                       ExpS_1.etq, ExpS_1.etq)
               parchea(Term.cod, Term.listaf, Term.listav,
                       Term.etg, Term.etg)
              OpAd.op
  ExpS_0.listaf = ExpS_0.listav = []
ExpS ::= Term
  ExpS.listaf = Term.listaf
  ExpS.listav = Term.listav
```



```
Term ::= Term OpMul Fact
  Termo.cod = parchea(Term1.cod, Term1.listaf, Term1.listav,
                         Term1.etg, Term1.etg)
              parchea(Fact.cod, Fact.listaf, Fact.listav,
                       Fact.etg, Fact.etg) | OpMul.op
  Termo.listaf = Termo.listav = []
Term ::= Fact
  Term.listaf = Fact.listaf
  Term.listav = Fact.listav
Fact ::= not Fact
  Fact<sub>1</sub>.cod = parchea(Fact<sub>1</sub>.cod, Fact<sub>1</sub>.listaf, Fact<sub>1</sub>.listav,
                        Fact<sub>1</sub>.etq, Fact<sub>1</sub>.etq)
  Facto.listaf = Facto.listav = []
Fact ::= iden
  Fact.listaf = Fact.listav = []
Fact ::= número
  Fact.listaf = Fact.listav = []
Fact ::= ( Exp )
  Fact.listaf = Exp.listaf
  Fact.listav = Exp.listav
```



```
ExpS ::= ExpS or Term
  ExpS_0.cod = parchea(ExpS_1.cod, ExpS_1.listaf, [],
                          ExpS_1.etq + 2, ?)
                copia | ir-v(?) | desapila | |
                parchea(Term.cod, Term.listaf, [], Term.etq, ?)
  ExpS_0.listav = ExpS_1.listav ++ Term.listav ++ [ExpS_1.etq + 1]
  ExpS_0.listaf = []
Term ::= Term and Fact
   Term<sub>0</sub>.cod = parchea(Term<sub>1</sub>.cod, [], Term<sub>1</sub>.listav,
                           ?, Term<sub>1</sub>.etq + 2)
                  copia | ir-f(?) | desapila |
                 parchea(Fact.cod, [], Fact.listav, ?, Fact.etq)
   Term<sub>0</sub>.listf = Term<sub>1</sub>.listaf ++ Fact.listaf ++ [Term<sub>1</sub>.etg + 1]
   Term<sub>o</sub>.listav = []
```



- * [] y ? se usan en parchea cuando realmente sólo interesa parchear una lista
- * ++ es la concatenación de listas

☐ En los esquemas de traducción optimizados queda así:

```
global cod, etq;
IWhile ::= \{var \text{ etqaux1, etqaux2}\}
              while
              \{\text{etgaux1} \leftarrow \text{etg}\}
              Exp
              do
              {parchea(Exp.listav,Exp.listaf,etq,etq);
                emite(ir-f(?));
                etgaux2 \leftarrow etgi
                etq \leftarrow etq + 1
             {emite(ir-a(etqaux1));
               etq \leftarrow etq + 1;
               parchea(etqaux2,etq)}
```



* parchea(Iv, If, ev, ef) realiza parcheo masivo de Iv con ev y de If con ef

```
global cod, etq;
ExpS(in listavh0, listafh0; out listav0, listaf0) ::=
       {var etqaux}
       ExpS(in listavh1, listafh1; out listav1, listaf1)
       {parchea([], listaf1, ?, etq + 2)}
       or
       {emite(copia);
        emite(ir-v(?));
        emite(desapila);
        etgaux \leftarrow etg + 1;
        etqaux \leftarrow etq + 3;}
       Term(in listavh2, listafh2; out listav2, listaf2)
       {parchea([], listaf2, ?, etq);
        listav0 = listav1 ++ listav2 ++ [etqaux];
        listaf0 = [];}
```



Traducción con precálculo de expresiones constantes

- Aunque la tarea del traductor no es interpretar el código fuente, precalcular el valor de las expresiones constantes será una excepción que realizaremos para optimizar el código fuente
- Esta optimización puede hacerse de varias formas
 - Realizando dos pasadas
 - 1. Transformando el código fuente
 - 2. Traduciendo el código fuente transformado
 - Realizando una pasada de transformación y traducción
 - □ El código fuente se transforma según se traduce (se complica un poco el proceso, pero será nuestro enfoque)
 - Realizando una pasada y trabajando sobre una representación intermedia del programa
 - ☐ A partir del código fuente se genera una *representación intermedia del programa* (= árbol de sintaxis abstracta)
 - ☐ Se transforma la representación intermedia del programa
 - Se genera código a partir de la representación intermedia del programa transformada



Traducción con precálculo de expresiones constantes

Reglas de optimización	
	Una expresión constante se ve reemplazada por su valor
	true y false son expresiones constantes de valor cierto y falso
	Número es expresión constante de valor valor (Número.lex)
	Cualquier operación Op sobre expresiones constantes X e Y es expresión constante de valor opera(valor(X), valor(Y))
	La asignación de una expresión constante X a una variable Id añade Id.lex a una tabla de variables constantes varscte (formada por pares lexema-valor) con el valor asociado valor (X)
	La asignación de una expresión no constante a una variable Id elimina Id de la tabla de variables constantes varscte
	Una variable Id de la <i>tabla de variables constantes</i> varscte es expresión constante de valor varscte.valor (Id.lex)
	Una instrucción if-then-else con expresión constante como condición es equivalente a su bloque then si la condición vale cierto, y a su bloque else si vale falso
	Una instrucción while-do con expresión constante como condición es equivalente a λ
	Eto



```
Proq ::= Decs & Is
 Prog.cod = Decs.cod ++ "&" ++ Is.cod
 Is.varscteh = []
 Prog.varscte = Is.varscte
Decs ::= Decs ; Dec
 Decso.cod = Decs1.cod ++ ";" ++ Dec.cod
Decs ::= Dec
 Decs.cod = Dec.cod
Dec ::= Tipo iden
 Dec.cod = Tipo.cod ++ " " ++ iden.lex
Tipo ::= bool
  Tipo.cod = "bool"
Tipo ::= int
  Tipo.cod = "int"
```



```
Is ::= Is ; I
  Iso.cod = si I.cod \neq \lambda
               Is1.cod ++ ";" ++ I.cod
            si no Isl.cod
  Is1.varscteh = Iso.varscteh
 I.varscteh = Isl.varscte
  Iso.varscte = I.varscte
Is ::= I
  Is.cod = I.cod
 I.varscteh = Is.varscteh
  Is.varscte = I.varscte
I ::= IAsiq
  I.cod = IAsiq.cod
  IAsiq.varscteh = I.varscteh
  I.varscte = IAsig.varscte
I ::= IIf
  I.cod = IIf.cod
  IIf.varscteh = IIf.varscteh
  IIf.varscte = IIf.varscte
```



```
I ::= IWhile
  I.cod = IIIf.cod
  TWhile.varscteh = TWhile.varscteh
  IWhile.varscte = IWhile.varscte
I ::= IComp
 I.cod = IComp.cod
  IComp.varscteh = I.varscteh
  I.varscte = IComp.varscte
IComp ::= begin IsOpc end
  IComp.cod = "begin "++IsOpc.cod++" end"
  IsOpc.varscteh = IComp.varscteh
  IComp.varscte = IsOps.varscte
IsOpc ::= \lambda
  IsOpc.cod = \lambda
  IsOpc.varscte = IsOpc.varscteh
IsOpc ::= Is
  IsOpc.cod = Is.cod
  Is.varscteh = IsOpc.varscteh
  IsOpc.varscte = Is.varscte
```



```
IAsiq ::= iden := Exp
  IAsig.cod = iden.lex ++ ":=" ++ si Exp.cte entonces Exp.val
                                  si no Exp.cod
  IAsiq.varscte = si Exp.cte entonces
                     añade(IAsig.varscteh,<iden.lex,Exp.val>)
                  si no elimina(IAsiq.varscteh,iden.lex)
 Exp.varscteh = IAsiq.varscte
IIf ::= if Exp then I PElse
  IIF.cod = si Exp.cte
             si Exp.val entonces I.cod
             si no PElse.cod
            si no "if "++Exp.cod++ " then "++I.cod++PElse.cod
  IIF.varscte = si Exp.cte
                 si Exp.val entonces I.varscte
                 si no PElse.varscte
                si no I.varscte ∩ PElse.varscte
  I.varscteh = IIf.varscteh
  PElse.varscteh = IIf.varscteh
```











```
Fact ::= not Fact
   Facto.cod = "not " ++ (si Fact1.cte entonces "pendiente"
                           si no Fact1.cod)
   Facto.cte = Fact1.cte
   Facto.val = eval("not", Fact1.val)
   Fact1.varscteh = Facto.varscteh
 Fact ::= ( Exp )
   Fact.cod = "("++ (si Exp.cte entonces Exp.val
                    si no Exp.cod) ++ ")"
   Fact.cte = Exp.cte
   Fact.val = Exp.val
   Exp.varscteh = Fact.varscteh
 Fact ::= num
   Fact.cte = true
   Fact.val = toNum(num.lex)
   Fact.cod = "pendiente 2"
```



```
ExpS ::= ExpS or Term
ExpSo.cod =
    si ExpS1.cte ^ ExpS1.val entonces "true"
    si no si ExpS1.cte ^ Term.cte entonces Term.cod
    si no si ExpS1.cte ^ Term.val entonces "true"
    si no si ExpS1.cte ^ Term.val entonces "false"
    si no si Term.cte ^ Term.val entonces "true"
    si no si Term.cte entonces ExpS1.cod
    si no ExpS1.cod ++ " and "++Term.cod
ExpSo.cte = (ExpS1.cte ^ ExpS1.val) \( \text{(Term.cte} ^ Term.val) \( \text{Verm.cte} ^ Term.val) \( \text{ExpS1.cte} ^ Term.cte \( \text{Term.cte} ^ Term.cte \( \text{Term.cte} ^ Term.cte \( \text{ExpS1.varscteh} = \text{ExpS0.varscteh} \)
```



```
Term ::= Term and Fact

Termo.cod =

si Term1.cte \ ¬Term1.val entonces "false"

si no si Term1.cte \ ¬Fact.cte entonces Fact.cod

si no si Term1.cte \ ¬Fact.val entonces "false"

si no si Term1.cte \ Fact.val entonces "true"

si no si Fact.cte \ ¬Fact.val entonces "false"

si no si Fact.cte entonces ExpS1.cod

si no Term1.cod ++ " and "++Fact.cod

Termo.cte = (Term1.cte \ ¬ Term1.val) \ \

(Fact.cte \ ¬ Fact.val) \ \

Term1.cte \ Fact.cte

Term1.varscteh = Fact.varscteh = Termo.varscteh
```



```
Fact ::= iden
    Fact.cte = <iden,_> ∈ Fact.varscteh
    Fact.val = valorDe(iden.lex, Fact.varscteh)
    Fact.cod = iden.lex
Fact ::= true
    Fact.cte = true
    Fact.val = true
    Fact.cod = true
Fact ::= false
    Fact.cte = true
    Fact.cte = true
    Fact.cte = true
    Fact.cte = true
    Fact.cod = "pendiente 3"
```



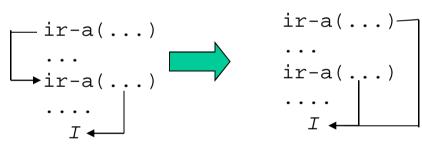
```
fun eval(op,v1,v2)
 si v1 = '?' v v2 = '?' entonces '?'
 si no si op = '+' entonces v1+v2
        si no si op = '*' entonces v1*v2
ffun
fun eval(op, v)
 si v = '?' entonces '?'
 si no si op = 'not'
           si v=0 entonces 1
           si no entonces 0
ffun
```



Traducción con sustitución de grupos de instrucciones

- A veces el código objeto generado contiene grupos de instrucciones inútiles o ineficientes que pueden eliminarse o sustituirse por otros (= hacer **optimización peephole**)
 - Existen situaciones triviales de tratar y otras más complejas
- Ejemplos:

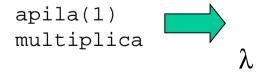
apila-dir(1)
$$\lambda$$
 desapila-dir(1)



$$ir-a(k)$$
 T
 $ir-a(k)$



(si *k está* fuera del grupo *l* y no hay ningún salto al grupo / en todo el programa)



Críticas, dudas, sugerencias...

Federico Peinado www.federicopeinado.es

