

Tema 9: Generación de Código Intermedio

1. Introducción
2. Código de tres direcciones
3. Generación de código intermedio mediante gramáticas de atributos
4. Estructuras y elementos de una matriz
5. Expresiones lógicas
6. Referencias no satisfechas. Relleno por retroceso
7. Instrucciones de control de flujo
8. Llamadas a subprogramas

1. Introducción

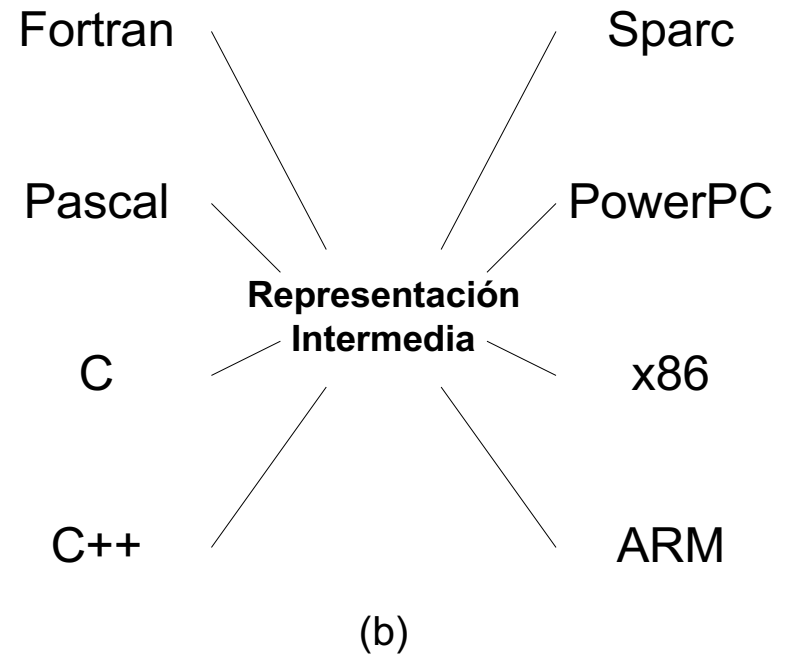
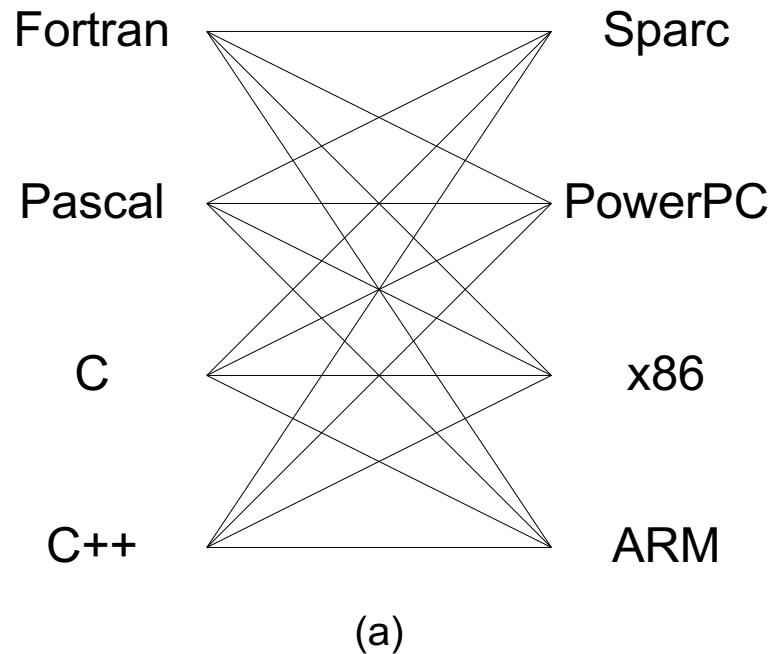
¿Qué lenguaje usar como destino del proceso de traducción?

- Otro lenguaje de alto nivel
- Código máquina
- **Representación intermedia**

Ventajas de usar representación intermedia

- Permite la división del proyecto en varias fases
- Facilita la reutilización de la etapa inicial (front-end) del compilador
- Se facilita la aplicación de transformaciones de mejora al código independientes de la arquitectura destino

Representaciones intermedias



Tipos de representaciones intermedia

- Grafos Dirigidos Acíclicos (GDA)
- Árbol de sintaxis abstracta (AST)
- Formato SSA (Static Single-Assignment)
- Código de tres direcciones

2. Código de 3 direcciones

Juego de instrucciones

Asignación	$x := y \text{ op } z$
Salto incondicional	goto E
Salto condicional	if $x \text{ op } y$ goto E
Pila de activación	push x $x := \text{pop}$
Llamada y	call E
Retorno de subprograma	ret
Asignaciones relativas	$a[i] := x$ $x := a[i]$
Acceso a la pila	FP TOP

3. Generación de CI mediante gramáticas atribuidas

Generación de CI para una asignación

Usaremos un procedimiento **emite** con un efecto colateral: Almacena código intermedio

Ej.

emite (x `:= ' 5 `+' 9)

SIGINST: Variable global con el número de la "SIGuiente INStrucción"

Asig \rightarrow id = E

E \rightarrow E + E

E \rightarrow id

Asig \rightarrow id = E	{ Asig.pos := BuscaPos (id.nom) ; emite (Asig.pos `:= ' E.pos); }
E \rightarrow E ₁ + E ₂	{ E.pos := <i>CrearVarTemp</i> () ; emite (E.pos `:= ' E ₁ .pos `+' E ₂ .pos) }
E \rightarrow id	{ E.pos := <i>BuscaPos</i> (id.nom) }

Árbol anotado: Ejemplo de generación de CI

Cadena: $a = b + c$

Nom	tipo	posición [nivel, desp]
a	tentero	[0,24]
b	tentero	[0,26]
c	tentero	[0,28]

Código tres direcciones generado

(usando posiciones de memoria)

(1) $[0,30] := [0,26] + [0,28]$

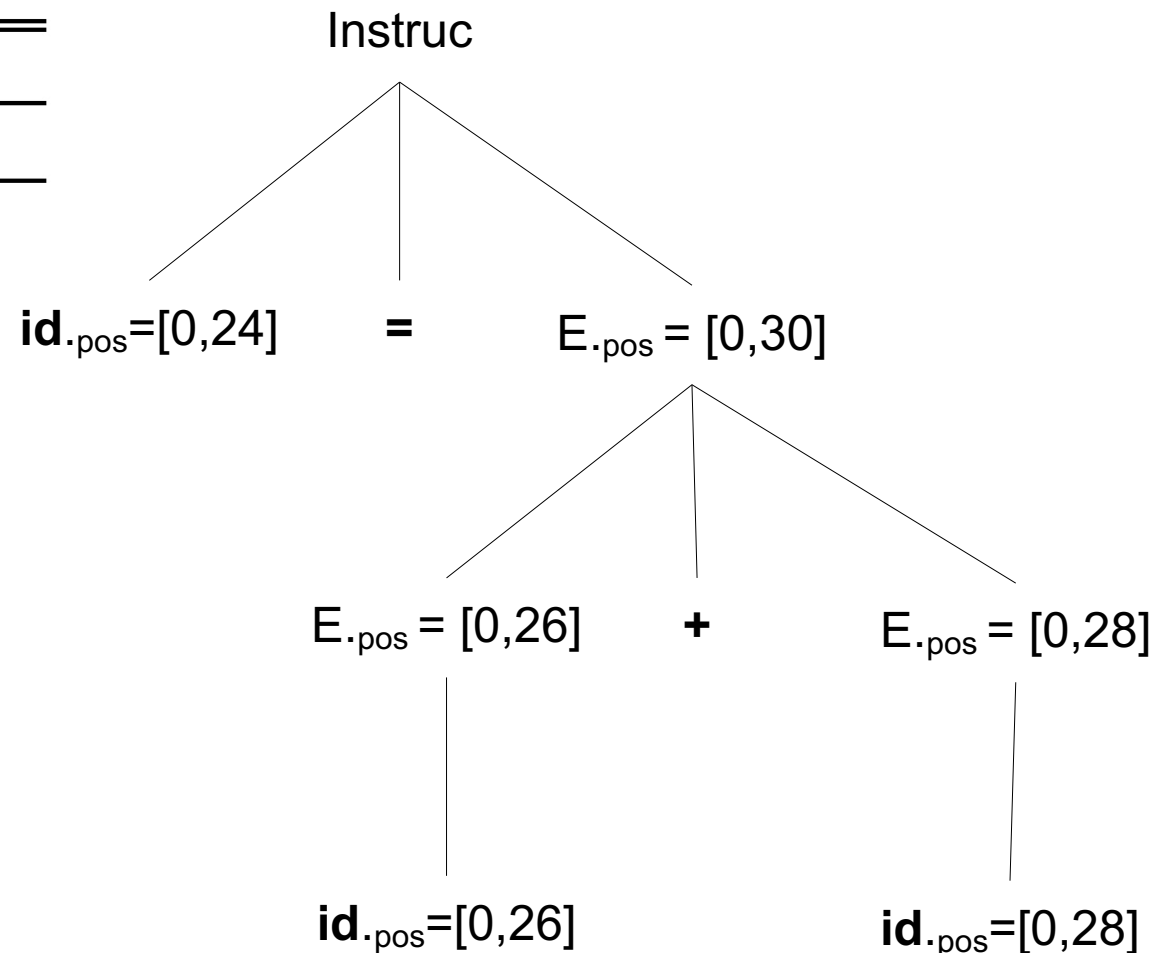
(2) $[0,24] := [0,30]$

Código tres direcciones generado

(usando nombres de posiciones de memoria)

(1) $t1 := b + c$

(2) $a := t1$



4. Estructuras y elementos de una matriz

Acceso a miembros de una estructura

```
struct ej {  
    int c1 ;  
    float c2 ;  
} a, b ;
```

TALLA_REAL = 4
TALLA_ENTERO = 2

TDS				
nom	tipo	posición	...	ref
a	testructura	0 , 124		
b	testructura	0 , 130		

Tabla de campos		
nom	tipo	posición
c1	tinteger	0
c2	treal	2



Acceso a miembros de una estructura. ETDS

$E \rightarrow id_1.id_2$	<pre>{ <u>si</u> BuscaTipo (id₁.nom) <> testructura <u>ent</u> MemError(); <u>sino</u> base := BuscaPos (id₁.nom) ; <u>si</u> no EsMiembro (id₁.nom, id₂.nom) <u>ent</u> MemError() <u>sino</u> desp_miembro:=BuscaPosMiembro(id₁.nom, id₂.nom); E.pos := base + desp_miembro ; }</pre>
$Asig \rightarrow id_1.id_2 = E$	<pre>{ <u>si</u> BuscaTipo (id₁.nom) <> testructura <u>ent</u> MemError(); <u>sino</u> base := BuscaPos (id₁.nom) ; <u>si</u> no EsMiembro (id₁.nom, id₂.nom) <u>ent</u> MemError() <u>sino</u> desp_miembro:=BuscaPosMiembro(id₁.nom, id₂.nom); Asig,pos := base + desp_miembro ; emite(Asig.pos `:=` E.pos); }</pre>

Acceso a elemento de una matriz

Declaración: $\text{int } A[n_1][n_2] \dots [n_n];$

Acceso al elemento $A[i_1][i_2] \dots [i_n]$: $\text{base} + (((i_1 * n_2 + i_2) * n_3 + i_3 \dots) * n_n + i_n) * \text{Talla}$
 n_i es el número de elementos de la i -ésima dimensión

IS \rightarrow id	{ LI.nom := id.nom }
LI = E	{ emite (LI.pos := LI.pos '*' Talla (id.nom)) ; IS.pos := BuscaPos (id.nom) ; emite (IS.pos '[' LI.pos ']' := E.pos ; }
LI \rightarrow [E]	{ LI.pos := CrearVarTemp() ; emite (LI.pos := E.pos); LI.ndim := 1; }
LI \rightarrow	{ LI ₁ .nom := LI.nom ; }
LI ₁ [E]	{ LI.ndim := LI ₁ .ndim + 1; LI.pos := LI ₁ .pos ; emite (LI.pos := LI.pos '*' Num_elementos(LI.nom, LI.ndim)) ; emite (LI.pos := LI.pos '+' E.pos) ; }

5. Expresiones lógicas

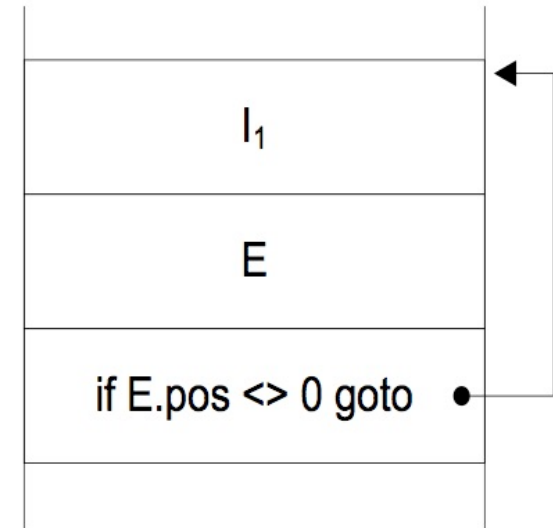
$E \rightarrow \text{true}$	$\{ E.\text{pos} := \text{CrearVarTemp}();$ $\text{emite}(E.\text{pos} \text{ ':=' } 1) \}$
$E \rightarrow \text{false}$	$\{ E.\text{pos} := \text{CrearVarTemp}();$ $\text{emite}(E.\text{pos} \text{ ':=' } 0) \}$
$E \rightarrow (E_1)$	$\{ E.\text{pos} := E_1.\text{pos} \}$
$E \rightarrow \text{id}$	$\{ E.\text{pos} := \text{BuscaPos}(\text{id}.\text{nom}) \}$

$E \rightarrow E_1 \text{ oprel } E_2$	$\{ E.\text{pos} := \text{CrearVarTemp}();$ $\text{emite}(E.\text{pos} \text{ ':=' } 1);$ $\text{emite}(\text{'if' } E_1.\text{pos oprel.op } E_2.\text{pos 'goto' SIGINST+2});$ $\text{emite}(E.\text{pos} \text{ ':=' } 0) \}$
--	---

$E \rightarrow E1 \text{ or } E2$	<pre> { E.pos:= CrearVarTemp(); emit(E.pos ':=' E1.pos '+' E2.pos) emit('if' E.pos '<= 1' goto' SIGINST+2); emit(E.pos ':=' 1') } </pre>
$ \ E1 \text{ and } E2$	<pre> { E.pos:= CrearVarTemp(); emit(E.pos ':=' E1.pos '*' E2.pos) } </pre>
$ \text{ not } E1$	<pre> { E.pos:= CrearVarTemp(); emit(E.pos ':=' 0); emit('if' E1.pos '<> 0 goto' SIGINST+2); emit(E.pos ':=' 1) } </pre>

Instrucciones de control de flujo: do-while


$I \rightarrow \text{do } I_1 \text{ while } (E)$



$I \rightarrow \text{do}$	{ I.inicio := SIGINST }
I_1	
$\text{while } (E)$	{ <i>emite</i> ('if' E.pos '<>' o goto' I.inicio) }

Instrucciones de control de flujo: while: *No funciona*

$I \rightarrow \text{while} (E) I$

$I \rightarrow \text{while}$ (E) I_1	<pre>{ I.inicio:= SIGINST } { emite('if' E.pos '=o goto' I.fin) } { emite('goto' I.inicio) ; I.fin := SIGINST }</pre> 
--	---

Esto no funciona. ¿Qué podemos hacer?

6. Relleno por retroceso

Perfil de las funciones usadas

ptro CreaLans(E)

Crea una lista que solo contiene un número de instrucción E a rellenar posteriormente. Devuelve un puntero a dicha lista.

void CompletaLans(ptro, E)

Rellena todas las instrucciones incompletas, cuyo número está contenido en la lista apuntada por ptro, con el valor de argumento E.

ptro FusionaLans (ptro, ptro)

Concatena las listas apuntadas por sus dos argumentos y devuelve un puntero a la nueva lista.

7. Instrucciones de control de flujo

while y if-else (usando relleno por retroceso)

$I \rightarrow \text{while}$ (E) I_1	$\{ I.inicio := SIGINST \}$ $\{ I.final := CreaLans(SIGINST);$ $\quad emite('if' E.pos '=o goto' ---) \}$ $\{ emite('goto' I.inicio);$ $\quad CompletaLans(I.final, SIGINST) \}$
--	--

$I \rightarrow \text{if } E$ $I_1 \text{ else}$ I_2	$\{ I.falso := CreaLans(SIGINST);$ $\quad emite('if' E.pos '=o goto' ---); \}$ $\{ I.fin := CreaLans(SIGINST);$ $\quad emite('goto' ---);$ $\quad CompletaLans(I.falso, SIGINST) \}$ $\{ CompletaLans(I.fin, SIGINST) \}$
---	--

For (usando relleno por retroceso)

$I \rightarrow \text{for} (I_1 ;$ $E ;$	$\{ I.\text{cond} := \text{SIGINST} ; \}$ $\{ I.\text{fin} := \text{CreaLans}(\text{SIGINST}) ;$ $\text{emite}(\text{'if' } E.\text{pos '}=0 \text{ goto' ---}) ;$ $I.\text{cuerpo} := \text{CreaLans}(\text{SIGINST}) ;$ $\text{emite}(\text{'goto' ---}) ;$ $I.\text{incr} := \text{SIGINST} ; \}$
$I_2)$	$\{ \text{emite}(\text{'goto' } I.\text{cond}) ;$ $\text{CompletaLans}(I.\text{cuerpo}, \text{SIGINST}) ; \}$
I_3	$\{ \text{emite}(\text{'goto' } I.\text{incr}) ;$ $\text{CompletaLans}(I.\text{fin}, \text{SIGINST}) ;$ $\}$

8. Llamadas a subprogramas

Decl_Subprg → Tipo id (Param_Form)	<pre> { emite('push FP'); emite('FP := TOP'); area_datos := CreaLans(SIGINST); emite('TOP := TOP + ---'); } </pre>
Bloque ;	<pre> { CompletaLans (area_datos, DESP) emite('TOP := FP'); emite('FP := pop'); emite('ret') } </pre>
E → id (Param_Act)	<pre> { E.pos := CreaVarTemp(); emite('TOP := TOP +' Talla_Tipo(id.nom)) } { emite('push' <estado máquina>); emite('call' BuscaPos(id.nom)); emite('pop' <estado máquina>); emite('TOP := TOP -' TallaParam(id.nom)); emite('pop' E.pos) } </pre>
Param_Act → E	<pre> { emite('push' E.pos) } </pre>
Param_Act → E , Param_Act	<pre> { emite('push' E.pos) } </pre>

ETDS de generación de C3D para:

$I \rightarrow \text{do } \{ I_1 \} \text{ for } (id ; E ; E_2)$

do-for es una instrucción repetitiva. Lo primero que hace la instrucción es inicializar la variable **id** a 0. Después de ello, mientras la expresión **E** sea cierta, se ejecutará el código de **I₁**, después el código de **I₂** y de nuevo se volverá a evaluar la expresión **E**. El bucle finalizará cuando **E** sea falso.

Ejercicio 2

Dada la siguiente gramática (inspirada en la proposición “switch” del C):

$I \rightarrow \text{switch} (E) \{ L \}$

$L \rightarrow \text{case num} : P ; L$

 | **default** : P ;

 | ε

...

$P \rightarrow P ; P$

 | **break**

 | I

 | ε

...

Donde:

Tanto “E” como la constante “num” deben ser enteras.

Si la constante de un “case” coincide con el valor de la expresión E, se deberá comenzar a ejecutar todas las instrucciones asociadas a ese “case” y a los que le siguen.

La proposición “break” provoca una salida inmediata del “switch”.

Las instrucciones asociadas al “default” se ejecutarán siempre (excepto si se encontró antes una instrucción “break”).

Construir un ETDS que realice las acciones semánticas necesarias para la generación de código intermedio de tres direcciones.