Tema 8: Representación de información en memoria

1. La tabla de símbolos (TDS)

- 1. Requisitos de la TDS
- 2. Implementación de la TDS
- 3. La TDS en un lenguaje con estructura de bloques

2. Gestión de memoria en tiempo de ejecución

- 1. Conceptos básicos
- 2. Asignación estática
- 3. Gestión del montículo (heap)
- 4. Gestión de la pila (stack) . Registros de activación
- 3. Ejemplo de asignación de memoria

1. La tabla de símbolos (TDS)

La Tabla de Símbolos

- La TDS es una estructra de datos que usa el PdL para almacenar información sobre los **símbolos** que aparecen en el **lenguaje fuente**.
- Durante las distintas fases de un compilador se almacena o lee información de la TDS.
- La información que contiene depende del propósito del PdL.
- Ej. Compilador: nombre de variable, tipo, posición de memoria,...

1.1. Requisitos de la TDS

- Operaciones básicas: Inserción, búsqueda y borrado.
- Ejemplo de información a incluir:

Tdsímbolos	Una entrada por cada objeto definido por el usuario.						
	Nombre	Lexema					
	Objeto	Categoría del objeto: <i>variable, parámetro, función</i>					
	Tipo	Tipo del objeto: <i>tinteger, tarray, trecord, tvacio</i> o <i>terror</i>					
	Des	Desplazamiento relativo en el segmento de datos					
	Niv	Nivel de anidamiento del objeto.					
	ref	Referencia a la tabla auxiliar: TdVectores , TdRegistros TdDominios					

TdVectores

Una entrada para cada dimensión del array, y el tipo de los elementos del array. Cada entrada contiene el número de elementos de la dimensión (o el límite inferior (min) y superior (max) según el lenguaje).

TdRegistros

Una entrada por cada campo de los registros.

Nombre Lexema

Tipo del campo. Tipo

Desplazamiento relativo del campo en el Des segmento correspondiente.

TdDominios Una entrada por cada dominio definido en los bloques.

Tratamiento de palabras reservadas:

- Por el analizador léxico: Un token específico para cada palabra reservada
- Como identificadores y tabla de palabras reservadas para diferenciar si se trata de una palabra reservada o de un identificados

Almacenamiento de los lexemas:

- Tabla de lexemas:
 - Mejor aprovechamiento de memoria
 - Elimina límite en la longitud de los lexemas

1.2. Implementación de la TDS

- Array de registros
- Listas enlazada ordenadas. Listas ordenadas doblemente enlazadas
- Árboles equilibrados ordenados
- Tablas de dispersión (hash)

1.3. TDS en lenguaje con estructura de bloques (LEB)

```
int a, b;
void function f1 (int p1, p2){
     int c, d;
     int function f2 (int p3, p4, p5){
           int e, c;
            ... }
     void function f3 (int p6, p7){
           f2(3,4,5); }
     f3(6,7); ... }
int function f4 {
     int f;
void function principal {
    f1(1,2); }
```

Problemas a resolver:

- Control del alcance de cada declaración
- Varios objetos con el mismo nombre accesibles

Posibles soluciones:

- a) Una subtabla para cada ámbito de declaraciones. Solución adecuada para compiladores de varias pasadas.
- b) Gestión como pila apoyada en una tabla de ámbitos. Solución no adecuada para compiladores de varias pasadas.

a) Gestión con una subtabla de símbolos por bloque

- Para cada *bloque* con declaraciones (por cada **ámbito** de declaraciones) se prepara una **subtabla** de símbolos.
- Cada subtabla de símbolos tendrá una entrada para cada símbolo declarado en el bloque.
- Las subtabla estarán **enlazadas** para representar el **anidamiento** de bloques.
- La búsqueda de símbolos en la TDS comienza por la subtabla del bloque más interior y continuará por la subtabla en la que ésta está definida.

```
var
                                                                     tentero
int a, b;
                                                                     tentero
                                                                var
void function f1 (int p1, p2){
                                                           f1
                                                               fun
                                                                     tvacio
   int c, d;
   if (p1 != 0) {
       int c, e;
                                                      tentero
                                            p1
                                                 par
                                            p2
                                                      tentero
                                                 par
   if (p2 > 3) {
       int d;
                                            C
                                                      tentero
                                                 var
                                                      tentero
                                                 var
                                            tentero
                                                           d
                                                                     tentero
                                       var
                                  C
                                                               var
                                            tentero
                                  e
                                       var
```

b) Gestión LIFO de TDS en LEB

- Al comenzar a compilar cada bloque iniciamos un nuevo ámbito. Se irán insertando en la TDS sus objetos locales.
- Al finalizar la compilación de un bloque se sacarán de la TDS todos sus objetos locales.
- Con esta gestión se garantiza que al compilar un bloque están en la TDS todos los objetos accesibles desde él, y solo los objetos accesibles desde él.

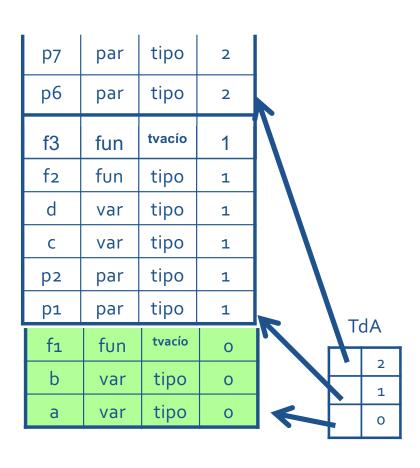
La Tabla de Ámbitos (TdA)

- Para hacer más eficiente la operación de desapilar todo los objetos locales a un bloque, podemos apoyarnos en una Tabla de Ámbitos (TdA).
- TdA: Estructura con una entrada por cada nivel de anidamiento de los bloques que aparecen en el programa fuente.
- TdA[n] apuntará al primer objeto definido en un bloque de nivel de anidamiento n insertado en la TDS.

Para sacar (desapilar) de la TDS objetos de nivel de anidamiento n: Top_TDS = TdA[n]

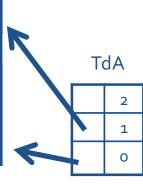
	l			ı
С	var	tipo	2	
е	var	tipo	2	
р5	par	tipo	2	
р4	par	tipo	2	
р3	par	tipo	2	
f2	fun	tipo	1	7
d	var	tipo	1	
С	var	tipo	1	
р2	par	tipo	1	
р1	par	tipo	1	TdA
f1	fun	tvacío	0	'\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \
b	var	tipo	0	1
а	var	tipo	0	0

```
int a, b;
void function f1 (int p1, p2){
    int c, d;
    int function f2 (int p3, p4, p5){
        int e, c;
        ... }
    void function f3 (int p6, p7){
        f2(3,4,5); }
    f3(6,7); ... }
int function f4 {
    int f;
        ... }
void function principal {
    f1(1,2); }
```



```
int a, b;
void function f1 (int p1, p2){
   int c, d;
   int function f2 (int p3, p4, p5){
     int e, c;
      ... }
   void function f3 (int p6, p7){
       f2(3,4,5); }
   f3(6,7); ... }
int function f4 {
   int f;
   ... }
void function principal {
  f1(1,2); }
```

f	var	tipo	1
f4	fun	tipo	0
f1	fun	tvacío	0
b	var	tipo	0
a	var	tipo	0



int a, b;
<pre>void function f1 (int p1, p2){</pre>
int c, d;
<pre>int function f2 (int p3, p4, p5){</pre>
int e, c ;
}
<pre>void function f3 (int p6, p7){</pre>
f2(3,4,5) ; }
f3(6,7); }
<pre>int function f4 {</pre>
<pre>int f ;</pre>
}
<pre>void function principal {</pre>
f1(1,2); }

2. Gestión de memoria en tiempo de ejecución

2.1. Conceptos básicos

Asignación dinámica:

• La localización del objeto solo se conocerá en tiempo de ejecución.

Asignación estática:

- La localización del objeto se conoce en tiempo de compilación.
- Requiere:
 - Conocer en tiempo de compilación el tamaño del objeto.
 - Conocer en tiempo de compilación el número de instancias simultaneas del objeto en ejecución.



2.1. Conceptos básicos

Ejemplo de esquema de memoria de programa en C

- Segmento de código (o segmento de "texto"): Contiene la instrucciones ejecutables
- Segmento de datos inicializados (o segmento de datos): Contienen variables globales y estáticas inicializadas

por el programador.

Puede dividirse en zona de solo lectura (para constantes) y de lectura-escritura para el resto.

- Segmento de datos sin inicializar (segmento .bss)
- Stack o pila de programa. La cima es apuntada por el stack pointer.
- Heap: Usado para asignar memoria dinámica (malloc,...)

Posiciones altas

Stack

Heap T

Datos globales sin inicializar

Datos globales inicializados

Segmento de código

2.2. Asignación estática

A los objetos se les asigna una dirección absoluta que se mantiene durante la ejecución del programa.

- Variable global
- Variables locales estáticas (que mantienen su valor entre llamadas a la función),
- Cadenas de caracteres,...
- Si el lenguaje no dispone de recursión, se puede emplear asignación estática para las variables locales.
 - Ej. Fortran 90.

Ejemplo de asignación estática

```
P \rightarrow
                                                                                        { NIVEL = 0 ; DESP = 0 ; }
                            L Decla
L Decla → Decla | L Decla
                                                                                                      Decla
Decla → DV
/**** Declaración de variables ****/
DV → T id; { InsertarTds (id.nom, "variable", T.tipo, NIVEL, DESP);
                                                                   DESP := DESP + T.talla :}
T \rightarrow int {T.tipo = Tentero; T.talla = TALLA ENTERO; }
         | float | { T.tipo = Treal; T.talla = TALLA REAL; }
         | bool | { T.tipo = Tlogico; T.talla = TALLA LOGICO; }
            | struct { C } { T.tipo := testructura (C.tipo); T.talla = C.talla }
C \rightarrow T id { C.tipo := (id.nom x T.tipo); C.talla = T.talla ;}
       C_1; T id C_2: C_3: C_4: C_4:
```

2.3. Gestión el montículo (heap)

Montículo: Región de memoria en la que los subbloques pueden ser asignados y liberados en cualquier orden.

- Se debe usar siempre que un objeto pueda cambiar de tamaño.
- Liberación de bloques:
 - Explícita: Indicada por el programador.
 - Implícita: El bloque asignado a un objeto debe liberarse automáticamente cuando se detecte que no se va a usar más el objeto. Requiere mecanismo recolector de basura (garbage collector) en tiempo de ejecución.
- Asignación de bloques: Problemas de fragmentación
 - Fragmentación interna:
 - El algoritmo de asignación asigna un bloque mayor del requerido para almacenar el objeto.
 - Fragmentación externa:
 - Los bloques asignados se van dispersando: Puede haber espacio disponible pero repartido en trozos tan pequeños que puede no ser suficiente para almacenar un objeto entero.

Asignación de memoria del montículo

- 1. Usando lista de bloques de memoria libres. Inicialmente hay un único bloque (todo el montículo).
- 2. Ante petición de memoria para objeto de tamaño t:
 - Algoritmo first-fit: Asigna primer bloque de la lista de tamaño >= t
 - Algoritmo best-fit: Asigna bloque más pequeño de la lista de tamaño >=t
- 3. El bloque asignado se divide en dos para asignar solo un tamaño t. El resto se mete en la lista de bloques libres.
- 4. Al liberar un bloque se comprueba si puede fusionarse con los bloques adyacentes (si están libres).

- Mejora: Mantener varias listas de bloques libres en función del tamaño de éstos.
- Para eliminar la fragmentación externa: Compactar (moviendo bloques asignados)

2.4. Gestión de la pila

Los bloques de memoria se asignan y liberan en orden LIFO

- Llamamos activación de un bloque a cada una de sus ejecuciones.
- Tiempo de vida de una activación es la **secuencia de pasos** entre el primer y último paso de la ejecución.
 - Los tiempos de vida de dos activaciones o no se solapan, o uno incluye completamente al otro.
- Árbol de activación: Representación del tiempo de vida de las activaciones de un programa:
 - Cada nodo representa una activación.
 - A es padre de B si B se activa durante el tiempo de vida de A
 - A está a la izquierda de B si el tiempo de vida de A es anterior a B (A finaliza antes de comenzar B)
- El flujo de control del programa coincide con un recorrido en profundidad del árbol de activación

Registro de activación

- Espacio de la pila de ejecución asignado a una activación para almacenar sus datos locales.
- En algunos lenguajes se llama marco (*frame*) de pila.
- Display: Bloque de punteros con una entrada por cada nivel de anidamiento de los objetos, donde Display[n] apunta al último RA con objetos de nivel de anidamiento n cargado. En algunos lenguajes es un solo registro

Estructura general

Variables locales y temporales

Enlaces de control

Parámetros

Ejemplo de RA y acceso (ej. Pascal)

+ TallaEstadoMaquina + TallaValorDevuelto)

```
Vbles. Temporales
                                                                                               ←TOP
Acceso a Vble. local (niv, desp):
vble.pos := display[niv] + desp
                                                                                    C
                                                              Display[2] \rightarrow
                                                                                    е
                                                                           Ptro. al RA del último
Acceso a Parámetro (niv, desp):
                                                                           bloque de iqual nivel
                                                                                 cargado
param.pos := display[niv] + desp
                                                                           Estado de la Máquina
  - (TallaTotalParametros + TallaEnlaces
                                                                                   p5
  + TallaEstadoMaquina )
                                                                                   p4
                                                                                   p3
                                                                              Valor devuelto
Depositar valor devuelto por una función de nombre (niv, -):
                                                                                 por pr2
 display[niv+1]
  - (TallaTotalParametros + TallaEnlaces
```

Ejemplo de Frame y acceso en C

• A los registros de activación se les suele denominar *Frames*

• No hay anidamiento de funciones: Todos los objetos serán globales o locales.

Display solo necesita 1 nivel -> FramePointer (FP)

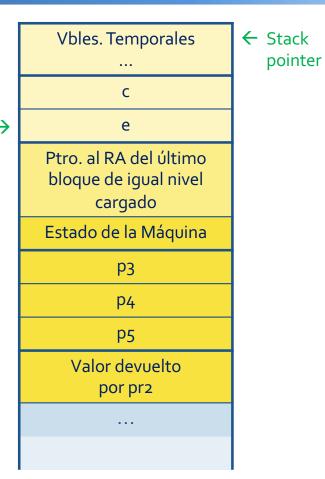
Nivel o = Global

Nivel 1 = Local (apuntado por FramePointer)

Ejemplo de Frame y acceso en C

```
Acceso a Vble. local (desp):
vble.pos := frame_pointer + desp
                                                  Frame Pointer >
Acceso a Parámetro (desp):
(apilados en orden inverso a declaración)
param.pos := frame_pointer
 - (desp + TallaEnlaces + TallaEstadoMaquina
        + TallaDelParametro)
Depositar valor devuelto por una función de nombre (
  niv, - ):
frame_pointer – ( TallaParámetros + TallaEnlaces
```

+ TallaEstadoMaquina + TallaValorDevuelto)



Secuencia de carga RA

Bloque llamador

- 1. Evaluar parámetros actuales
- 2. Si la función devuelve un valor: Crear variable temporal para el valor devuelto
- 3. Apilar parámetros actuales
- 4. Apilar Estado de la Máquina (incluida la dirección de retorno) y saltar al código del bloque llamado (CALL)

Bloque llamado

- Apilar enlace de control (Apilar Display[n])
- Actualizar display (display[n] = TOP)
- 3. Reservar espacio para área de datos locales

Secuencia de descarga RA

Bloque llamado

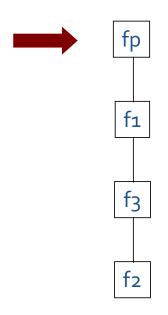
- Desapilar área de datos (TOP= Display[n])
- Desapilar enlace de control y restaurar valor de display (display[n]= POP)
- 3. Desapilar Estado de la máquina, incluida la dirección de retorno y saltar a ella (RETURN)

Bloque llamador

- 1. Desapilar parámetros actuales
- 2. Si la función devuelve un valor: Desapilar valor de retorno

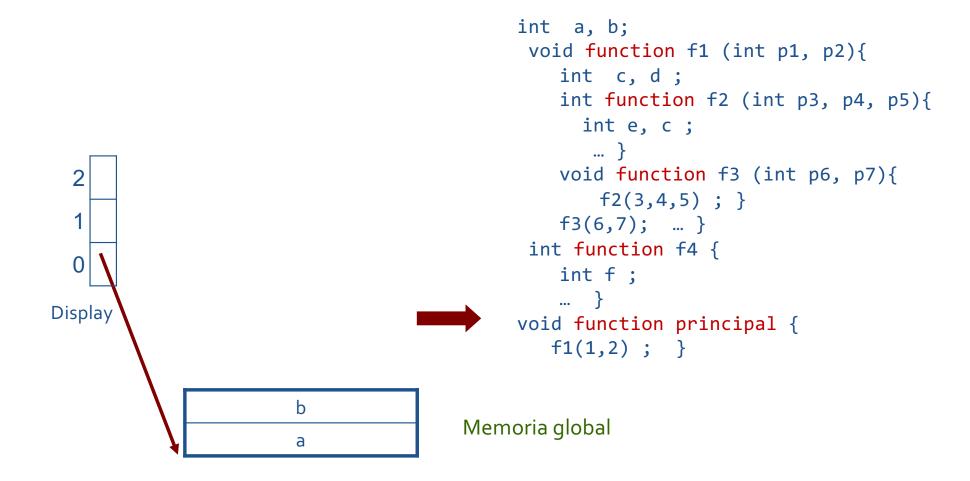
Ejemplo (I)

El árbol de activación es:

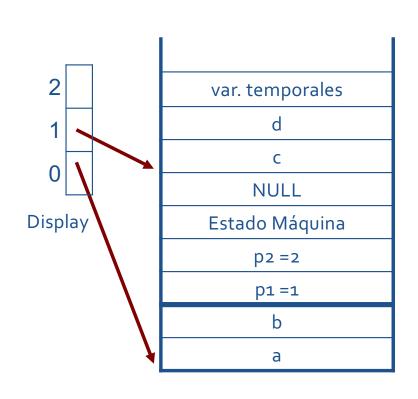


```
int a, b;
void function f1 (int p1, p2){
   int c, d;
   int function f2 (int p3, p4, p5){
      int e, c;
      ... }
   void function f3 (int p6, p7){
      f2(3,4,5); }
   f3(6,7); ... }
int function f4 {
   int f;
   ... }
void function principal {
   f1(1,2); }
```

Ejemplo (II)

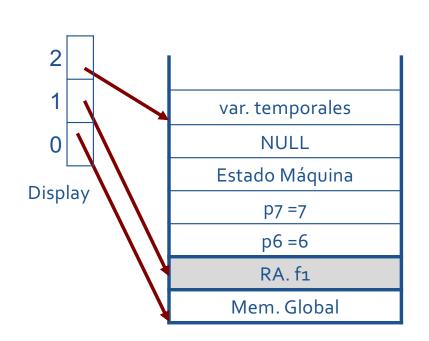


Ejemplo (III)



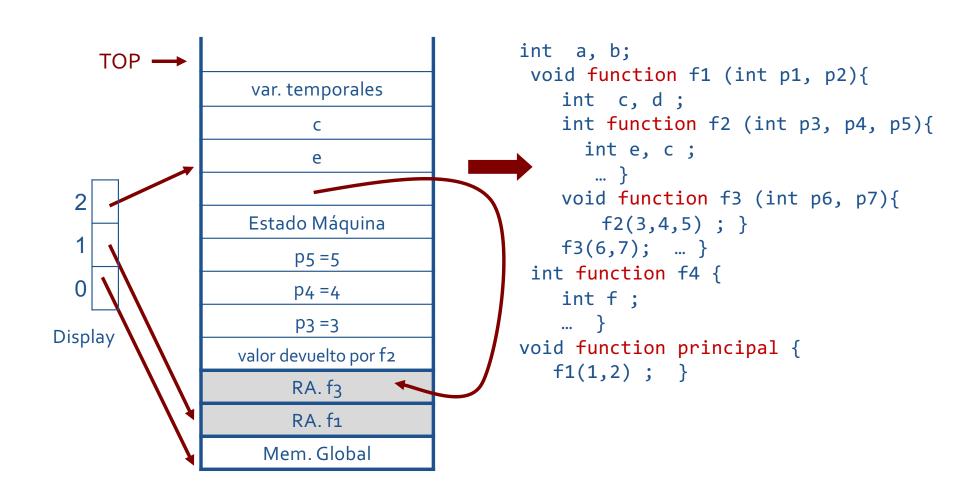
```
int a, b;
void function f1 (int p1, p2){
   int c, d;
   int function f2 (int p3, p4, p5){
     int e, c;
   void function f3 (int p6, p7){
       f2(3,4,5); }
   f3(6,7); ... }
 int function f4 {
   int f;
void function principal {
   f1(1,2); }
```

Ejemplo (IV)

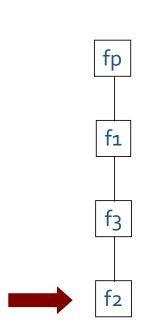


```
int a, b;
void function f1 (int p1, p2){
   int c, d;
   int function f2 (int p3, p4, p5){
     int e, c;
   void function f3 (int p6, p7){
       f2(3,4,5); }
   f3(6,7); ... }
 int function f4 {
   int f;
void function principal {
   f1(1,2); }
```

Ejemplo (V)



Ejemplo (VI)



```
int a, b;
void function f1 (int p1, p2){
    int c, d;
    int function f2 (int p3, p4, p5){
        int e, c;
        ... }

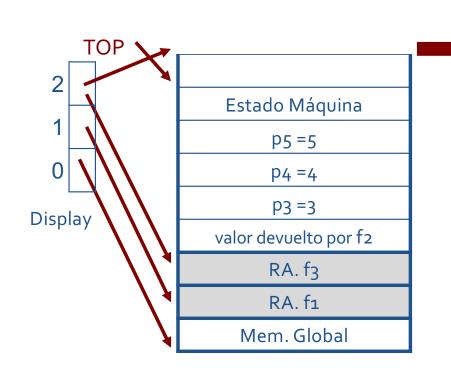
void function f3 (int p6, p7){
        f2(3,4,5); }

f3(6,7); ... }

int function f4 {
    int f;
        ... }

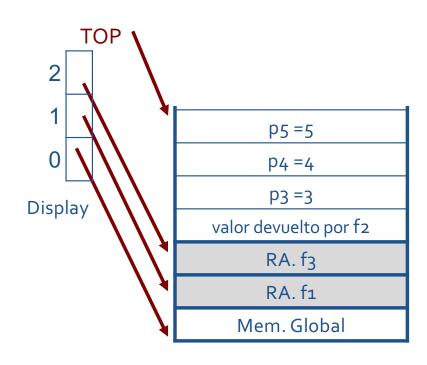
void function principal {
    f1(1,2); }
```

Ejemplo (VII)



```
int a, b;
void function f1 (int p1, p2){
   int c, d;
   int function f2 (int p3, p4, p5){
      int e, c;
      ... }
   void function f3 (int p6, p7){
      f2(3,4,5); }
   f3(6,7); ... }
int function f4 {
   int f;
   ... }
void function principal {
   f1(1,2); }
```

Ejemplo (VIII)



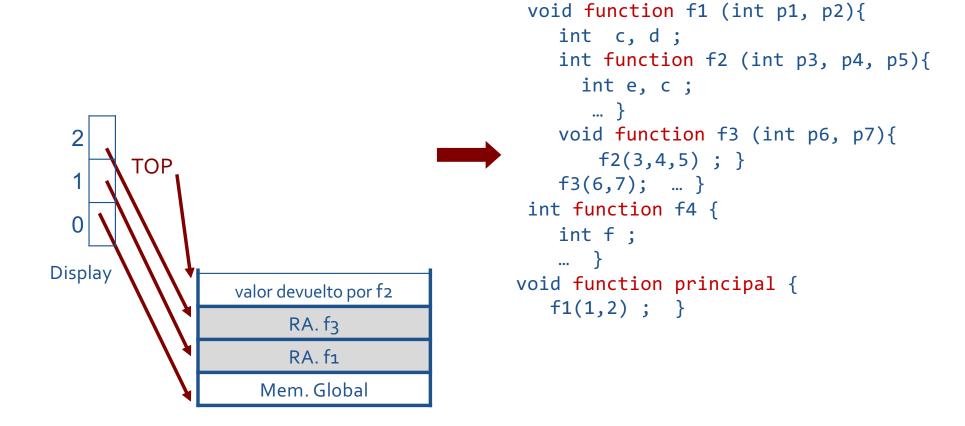
```
int a, b;
void function f1 (int p1, p2){
    int c, d;
    int function f2 (int p3, p4, p5){
        int e, c;
        ... }

void function f3 (int p6, p7){
        f2(3,4,5); }
        f3(6,7); ... }

int function f4 {
        int f;
        ... }

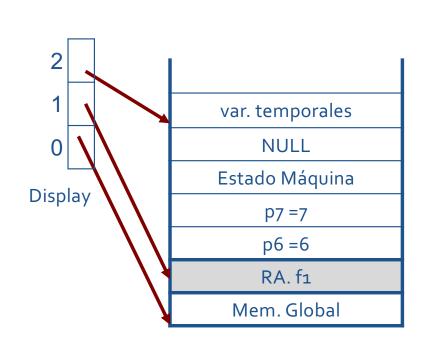
void function principal {
        f1(1,2); }
```

Ejemplo (IX)



int a, b;

Ejemplo (X)



```
int a, b;
void function f1 (int p1, p2){
   int c, d;
   int function f2 (int p3, p4, p5){
     int e, c;
   void function f3 (int p6, p7){
       f2(3,4,5); }
   f3(6,7); ... }
 int function f4 {
   int f;
void function principal {
   f1(1,2); }
```

Ejemplo de asignación de memoria

```
P \rightarrow L_Decla
L\_Decla \to Decla
         L_Decla Decla
Decla → DV
/**** Declaración de variables ****/
DV \rightarrow T id;
  T \rightarrow int
   | float
   | bool
   | struct {C}
/*** Miembros de estructuras ****/
C \rightarrow T id
    \mid C_1; T id
```

3. Asignación de memoria

Ejemplo de asignación de memoria (1/4)

```
P \rightarrow
       L Decla
L_Decla → Decla | L_Decla Decla
Decla \rightarrow DV | DF
/**** Declaración de variables ****/
DV \rightarrow T id; | T id LI;
T \rightarrow int
   | float
   bool
   | struct {C}
/**** Miembros de estructuras ****/
C \rightarrow T id
   \mid C_1; Tid
/**** Índices de declaración de array ****/
L_l \rightarrow [cte] \mid L_l, [cte]
```

```
/**** Declaración de función ****/
DF → Cab_F Bloque
Cab_F\rightarrowT id (P_F)
Bloque \rightarrow { DVL L_Inst }
DVL \rightarrow DVL DV \mid \epsilon
/**** Parámetros formales ****/
P F \rightarrow L PF \mid \epsilon
L_PF \rightarrow T id \mid Tid, L_PF_1
```

Ejemplo de asignación de memoria (2/4)

```
P \rightarrow
                          { NIVEL = 0 ; DESP = 0 ; }
        L_Decla
L_Decla → Decla | L_Decla Decla
Decla → DV | DF
/**** Declaración de variables ****/
DV → T id; { InsertarTds (id.nom, "variable", T.tipo, NIVEL, DESP);
                 DESP := DESP + T.talla ; }
DV \rightarrow T id L_I; {InsertarTds (id.nom, "variable", vector(T.tipo), NIVEL, DESP);
                DESP := DESP + (L I.talla * T.talla); }
T \rightarrow int
          {T.tipo = Tentero; T.talla = TALLA_ENTERO; }
  bool
         { T.tipo = Tlogico; T.talla = TALLA_LOGICO; }
   | struct { C } { T.tipo := testructura (C.tipo); T.talla = C.talla }
```

Ejemplo de asignación de memoria (3/4)

/**** Índices de declaración de array ****/

```
L_I \rightarrow [cte] \qquad \{LI.talla := cte.lexval;\}
|L_I_1[cte] \qquad \{L_I.talla := cte.lexval * L_I_1.talla;\}
```

```
C \rightarrow T id \{C.tipo := (id.nom \times T.tipo); C.talla = T.talla\}
\{C.tipo := (id.nom \times T.tipo); C.talla = C1.talla + T.talla\}
```

Ejemplo de asignación de memoria (4/4)

/**** Declaración de función ****/

/**** Parámetros formales ****/

```
 \begin{array}{ll} \mathsf{P}_{-}\mathsf{F} \to & \{ \mathit{L}_{-}\mathsf{PF}. \mathit{tallapar} := \mathit{TALLA}_{-}\mathit{EC} + \mathit{TALLA}_{-}\mathit{EM} \, ; \, \} \\ & \mathsf{L}_{-}\mathsf{PF} & \{ \mathit{P}_{-}\mathit{F}. \mathit{tipo} := \mathit{L}_{-}\mathit{PF}. \mathit{tipo} \, \} \\ & \mathsf{I} \quad & \{ \mathit{P}_{-}\mathit{F}. \mathit{tipo} := \mathit{tvacio} \, \} \\ & \mathsf{L}_{-}\mathsf{PF} \to \mathsf{T} \quad \mathsf{id} & \{ \mathit{InsertarTds} \, (\mathit{id}. \mathit{nom}, \mathit{"parametro"}, \mathit{T}. \mathit{tipo}, \, \mathit{NIVEL}, \, - \, (\mathit{L}_{-}\mathit{PF}. \mathit{tallapar} + \mathit{T}. \mathit{talla}) ) \, ; \\ & \mathsf{L}_{-}\mathit{PF}. \mathit{tipo} := \mathit{T}. \mathit{tipo} \, ; \, \} \\ & \mathsf{I} \quad & \{ \mathit{L}_{-}\mathit{PF}_{1}. \mathit{tallapar} := \mathit{L}_{-}\mathit{PF}. \mathit{tallapar} + \mathit{T}. \mathit{talla} \, ; \\ & \mathit{InsertarTds} \, (\mathit{id}. \mathit{nom}, \mathit{"parametro"}, \mathit{T}. \mathit{tipo}, \, \mathit{NIVEL}, \, - \, \mathit{L}_{-}\mathit{PF}_{1}. \mathit{tallapar}) \, \} \\ & \mathsf{L}_{-}\mathit{PF}_{1} & \{ \mathit{L}_{-}\mathit{PF}. \mathit{tipo} := \mathit{L}_{-}\mathit{PF}_{1}. \mathit{tipo} \times \mathit{T}. \mathit{tipo} \, \} \\ \end{array}
```

Ejercicio 1

Suponiendo que la talla de los enteros es 2, de los reales y enlaces de control 4, y del "estado de la máquina" 10:

- a) Muestra el contenido de la TDS en los puntos marcados como TDS
- b) Indica el desplazamiento relativo de i y x en el punto de control PC.
- c) Muestra el contenido de la pila de ejecución cada vez que se pasa por el punto PILA.

Ejercicio 2

```
float prod( float a, float b) {
return a*b }
float sum (int a, int b) {
  float t
 if (a \le 5) { t= prod(a,b); return t; } /* \leftarrow TDS
                                                 /* ← PILA */
  else {
  return suma(a-1, b-1);}
                                                  /* ← PC1 */
int main () {
  int a; int b;
 a = 7;
                                    /* ← PC2 */
                                    /* ← TDS */
  b = 9;
  printf( "Resultado %f\n", sum(a, b)); }
```

- a) Muestra contenido de TDS en puntos TDS (supón que la talla de enteros es 2 y de reales 4)
- b) Indica el desplazamiento relativo de a y b en los puntos de control: PC1 y PC2.
- c) Muestra contenido de la pila de ejecución cada vez que se pasa por el punto PILA.