Tema 2. Instruccions i tipus de dades bàsics

Joan Manuel Parcerisa





Instruccions i tipus de dades bàsics

- Introducció al MIPS
- Operands: en registre, immediats, o en memòria
- Sistemes de representació binària
 - Naturals i Enters (repàs)
 - Caràcters
 - Instruccions MIPS
- Vectors i punters

Introducció: RISC vs CISC

- CISC (Complex Instruction Set Computer)
 - Joc d'instruccions gran i complex
 - Instruccions de mida i nombre d'operands variable
 - Múltiples formats d'instrucció i modes d'adreçament
 - Difícil d'optimitzar les operacions més simples i frequents
 - Exemple: x86
- RISC (Reduced Instruction Set Computer)
 - Poques instruccions
 - Mida fixa i pocs operands
 - Formats d'instrucció i modes d'adreçament simples
 - Lema: "make the common case fast"
 - Exemple: MIPS, ARM, RISC-V, PowerPC...
- Actualment s'han difuminat les diferències
 - Les CPU x86 (PentiumPro, 1995) descomponen internament les instruccions complexes en microoperacions tipus-RISC
 - Els repertoris dels RISC han crescut igual que els CISC

El processador MIPS

- Microprocessor without Interlocked Pipeline Stages
 - Dissenyat per J.L. Hennessy a Stanford, 1981
 - Tipus RISC, usat en múltiples dispositius
 - Routers, Impressores, Playstation (PS2, PSP)
 - Àmpliament usat en docència
- Successives versions del ISA de MIPS
 - En EC estudiarem la versió MIPS32

EI ISA MIPS32

• MIPS32

- 32 registres de 32 bits cadascun
- Paraules (words) de 32 bits (4 bytes)
 - Normalment word = "dada de la mida dels registres"
- Instruccions i adreces de 32 bits
- Memòria: 2³² bytes adreçables (1 byte per adreça)
- Com emmagatzema en memòria números multi-byte?
 - Little-endian o big-endian, MIPS ho decideix durant l'arrancada
 - En EC assumirem sempre little-endian. Per exemple: guardar el word 0x76543210 a l'adreça 0x10010000

adreça	
0x10010000	0x10
0x10010001	0x32
0x10010002	0x54
0x10010003	0x76

Declaració de variables

En assemblador MIPS

- NO hi ha declaracions de variables
- Hi ha etiquetes (que identifiquen adreces de memòria)
- Hi ha directives que reserven X bytes, i els inicialitzen

En C

- Hi ha declaracions de variables, on s'especifica el seu tipus
- La mida en bytes de cada tipus depèn de l'ISA on s'implementa
- En EC assumirem les següents mides:

Tipus de variable en C	Directiva de reserva d'espai en MIPS	Mida en bytes
char / unsigned char	.byte	1
short / unsigned short	.half	2
int / unsigned int	.word	4
long long / unsigned long long	.dword	8

Quan les operem, assumim que són naturals

Quan les operem, assumim que són enters en Ca2

Variables globals i locals

Exemple en C

```
int g1, g2;
void main() {
    int loc1, loc2;
    loc1 = g1;
    ...
}
```

- Tipus de variables en C: auto, extern, static, register
- Per defecte, les declarades FORA d'una funció són extern
 - Accessibles des de qualsevol funció
 - Ocupen una adreça fixa de memòria durant tot el programa
 - En EC les anomenarem globals
- Per defecte, les declarades DINS una funció són auto
 - Accessibles sols dins el bloc on es declaren
 - Ocupen memòria (o registres) sols mentre s'executa la funció
 - En EC les anomenarem locals

Declaració de variables. Exemple

Codi en C

Codi assemblador MIPS

```
int g = 0x76543210;

void func() {
    int loc1;

    loc1 = 13;
    ...
}

addiu $t0,$zero,13 #loc1=13
...
```

GLOBAL: g ocuparà l'adreça fixa 0x10010000 (sinònim de l'etiqueta g) durant tot el programa

LOCAL: **loc1** ocupa el registre \$t0 mentre s'executa *func*. Després, \$t0 pot allotjar altres dades

Declaració de variables. Més exemples

• Declaracions globals en C...

Declaració de variables. Més exemples

• Declaracions globals en C...

· ... i la seva traducció a assemblador

```
.data
c: .byte 0x11
s: .half 0x2211
i: .word 0x44332211
iext: .word 0xFF
iz: .word 0
ui: .word 0
l: .dword 0x8877665544332211
```

Alineació automàtica

Les variables en memòria han d'estar alineades:

L'adreça de la variable ha de ser múltiple de la seva mida

Altrament es poden produir excepcions durant l'execució

Tinus do variable en C	Directiva de reserva	Mida en	Adreça
Tipus de variable en C	d'espai en MIPS	bytes	múltiple de
char / unsigned char	.byte	1	-
short / unsigned short	.half	2	2
int / unsigned int	.word	4	4
long long / unsigned long long	.dword	8	8

o Les directives .half, .word, .dword ja alineen automàticament

Alineació automàtica. Exemple

• En C

```
unsigned char a;
short b = 13;
char c = -1, d= 10;
int e = 0x10AA00FF;
long long f = 0x7766554433221100;
```

Assemblador

```
.data
a: .byte 0
b: .half 13
c: .byte -1
d: .byte 10
e: .word 0x10AA00FF
f: .dword 0x7766554433221100
```

Alineació automàtica. Exemple

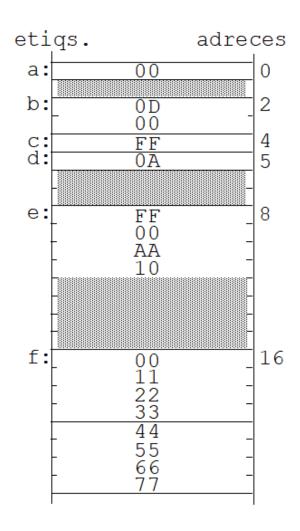
• En C

```
unsigned char a;
short b = 13;
char c = -1, d= 10;
int e = 0x10AA00FF;
long long f = 0x7766554433221100;
```

Assemblador

```
.data
a: .byte 0
b: .half 13
c: .byte -1
d: .byte 10
e: .word 0x10AA00FF
f: .dword 0x7766554433221100
```

Memòria



Alineació explícita amb .align

Vector inicialitzat

○ La directiva .half assegura l'alineació correcta de vec ✓

Alineació explícita amb .align

Vector inicialitzat

- La directiva .half assegura l'alineació correcta de vec ✓
- Vector global sense inicialitzar (zeros, per defecte)

- La directiva .space n reserva n bytes i els posa a 0
- Però no fa cap alineació: vec pot quedar mal alineat!

Alineació explícita amb .align

Vector inicialitzat

- La directiva .half assegura l'alineació correcta de vec ✓
- Vector global sense inicialitzar (zeros, per defecte)

- La directiva .space n reserva n bytes i els posa a 0
- Però no fa cap alineació: vec pot quedar mal alineat!
- Cal forçar l'alineació amb .align n, essent n ∈ [1, 2, 3]
 - Ubica l'etiqueta vec a la primera adreça disponible múltiple de 2ⁿ

Operands

- En mode registre
- En mode immediat
- En memòria

Operands

- Mode d'adreçament és la manera d'especificar l'operand
 - MIPS suporta 5 modes
 - o registre, immediat, memòria, pseudodirecte i relatiu al PC
- Operands en mode registre
 - L'operand resideix en un registre
 - La instrucció especifica l'identificador del registre (5 bits)
 - Suma i resta:

```
addu rd, rs, rt \# rd \leftarrow rs + rt subu rd, rs, rt \# rd \leftarrow rs - rt
```

Registres

Número	Nom	Utilització
\$0	\$zero	Sempre val zero, no modificable
\$1	\$at	Reservat a l'expansió de macros (convé no usar-lo)
\$2-\$3	\$v0-\$v1	Resultat de subrutines (sols usarem \$v0)
\$4-\$7	\$a0-\$a3	Arguments o paràmetres de subrutines
\$8-\$15	\$t0-\$t7	Temporals
\$16-\$23	\$s0-\$s7	Saved ("segurs"), es preserven en cridar a una subrutina
\$24-\$25	\$t8-\$t9	Temporals
\$26-\$27	\$k0-\$k1	Reservats per al nucli (Kernel) del SO (convé no usar-los)
\$28	\$gp	Global pointer, explicat al Tema 3 (no l'usarem)
\$29	\$sp	Stack pointer, conté l'adreça del cim de la pila
\$30	\$fp	Frame pointer (no l'estudiem)
\$31	\$ra	Return address, adreça de retorn de subrutina

Operands en mode immediat

- L'operand de 16 bits es codifica en la pròpia instrucció
- Quan s'executa, converteix l'operand de 16 a 32 bits
 - Per extensió de signe (enters en Ca2)
 - Per extensió de zeros (operands sense signe)

Alguns exemples:

```
addiu rt, rs, imm16 # rt \leftarrow rs + SignExt(imm16)

ori rt, rs, imm16 # rt \leftarrow rs OR ZeroExt(imm16)

lui rt, imm16 # rt<sub>31..16</sub> \leftarrow imm16

# rt<sub>15..0</sub> \leftarrow 0x0000
```

Donat el codi en C:

```
int f, g, h, i;
...
f = (g + h) - (i - 100);
```

- Suposem que f, g, h, i estan als registres \$t0, \$t1, \$t2, \$t3
- Traduir la sentència a assemblador MIPS

```
addu $t4, $t1, $t2  # g+h

addiu $t5, $t3, -100  # i-100

subu $t0, $t4, $t5  # f = (...) - (...)
```

Operands en mode memòria

- Sols s'admeten en instruccions de tipus load i store
 - → Cal carregar les dades de memòria en registres per usar-les en operacions aritmètico-lògiques
- Lectura d'un word en memòria: Load Word

```
lw rt, off16(rs) # rt \leftarrow Mem<sub>word</sub>[rs + SignExt(off16)]
```

Escriptura d'un word en memòria: Store Word

```
sw rt, off16(rs) # Mem_{word}[rs + SignExt(off16)] \leftarrow rt
```

Donat el codi en C:

```
int A[100], h;
...
A[12] = h + A[8];
```

- Suposem que h està en \$t2, i l'adreça base de A en \$t3
- Traduir la sentència a assemblador MIPS

```
lw $t0, 32($t3) # $t0 \leftarrow A[8]
addu $t0, $t2, $t0 # sumar h + A[8]
sw $t0, 48($t3) # A[12] \leftarrow $t0
```

Operands en memòria halfword amb signe

Load Halfword amb signe

```
    th rt, off16(rs) # rt ← SignExt[Memhalf[rs + SignExt(off16)]]
    Copia un halfword de memòria als 2 bytes de menor pes de rt
    El converteix a 4 bytes extenent el signe (replica el bit 15)
```

Store Halfword

```
sh rt, off16(rs) # Mem<sub>half</sub>[rs + SignExt(off16)] ← rt<sub>15...0</sub>
o Copia a memòria els 2 bytes de menor pes de rt
```

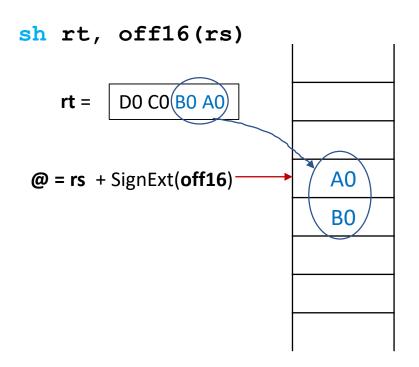
Operands en memòria halfword amb signe

Load Halfword amb signe

- 1h rt, off16(rs) # rt \leftarrow SignExt[Mem_{half}[rs + SignExt(off16)]]
 - Copia un halfword de memòria als 2 bytes de menor pes de rt
 - El converteix a 4 bytes extenent el signe (replica el bit 15)

Store Halfword

- sh rt, off16(rs) # $Mem_{half}[rs + SignExt(off16)] \leftarrow rt_{15..0}$
 - Copia a memòria els 2 bytes de menor pes de rt



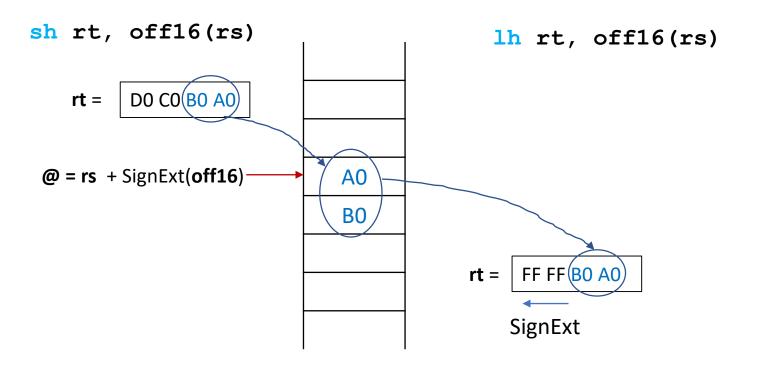
Operands en memòria halfword amb signe

Load Halfword amb signe

- lh rt, off16(rs) # rt SignExt[Mem_{balf}[rs + SignExt(off16)]]
 - Copia un halfword de memòria als 2 bytes de menor pes de rt
 - El converteix a 4 bytes extenent el signe (replica el bit 15)

Store Halfword

- sh rt, off16(rs) # $Mem_{half}[rs + SignExt(off16)] \leftarrow rt_{15..0}$
 - o Copia a memòria els 2 bytes de menor pes de rt



Operands en memòria byte amb signe

Load Byte, amb signe

```
    rt, off16(rs) # rt ← SignExt[Mem<sub>byte</sub>[rs + SignExt(off16)]]
    Copia un byte de memòria al byte de menor pes de rt
    El converteix a 4 bytes extenent el signe (replica el bit 7)
```

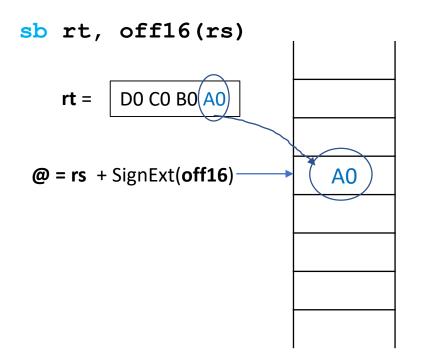
Store Byte

```
sb rt, off16(rs) # Mem<sub>byte</sub>[rs + SignExt(off16)] ← rt<sub>7..0</sub>
o Copia a memòria el byte de menor pes de rt
```

Operands en memòria byte amb signe

Load Byte, amb signe

- rt, off16(rs) # rt ← SignExt[Mem_{byte}[rs + SignExt(off16)]]
 Copia un byte de memòria al byte de menor pes de rt
 El converteix a 4 bytes extenent el signe (replica el bit 7)
- Store Byte
 - sb rt, off16(rs) # $Mem_{byte}[rs + SignExt(off16)] \leftarrow rt_{7..0}$
 - Copia a memòria el byte de menor pes de rt



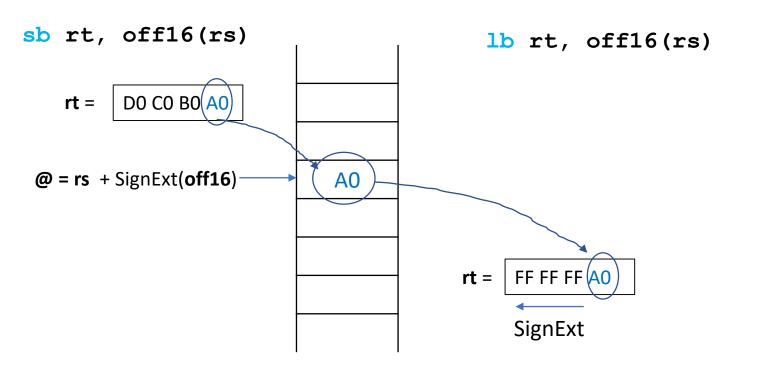
Operands en memòria byte amb signe

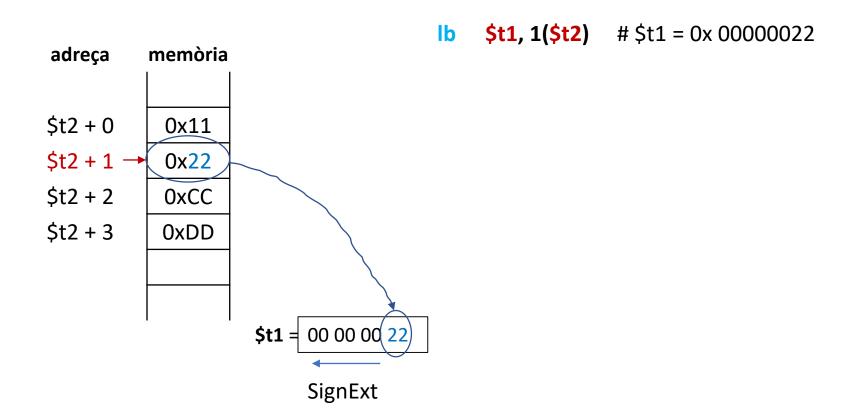
Load Byte, amb signe

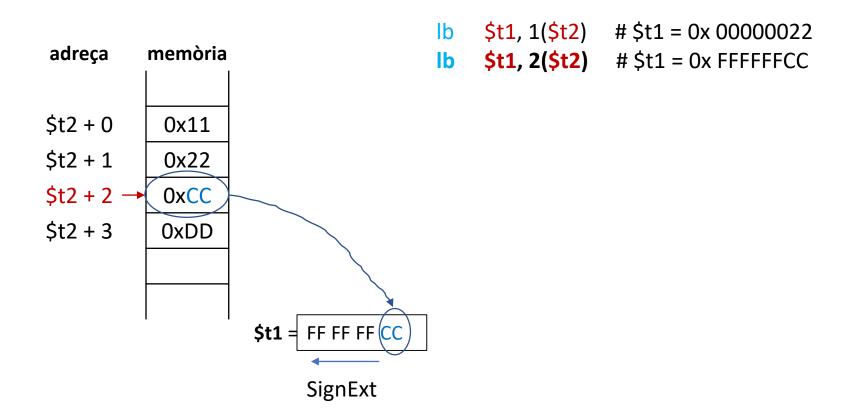
- tr off16(rs) # rt ← SignExt[Mem_{byte}[rs + SignExt(off16)]]
 Copia un byte de memòria al byte de menor pes de rt
 - El converteix a 4 bytes extenent el signe (replica el bit 7)

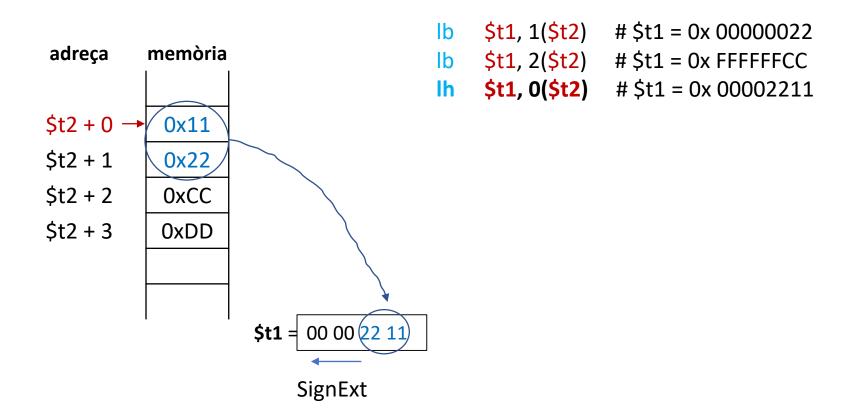
Store Byte

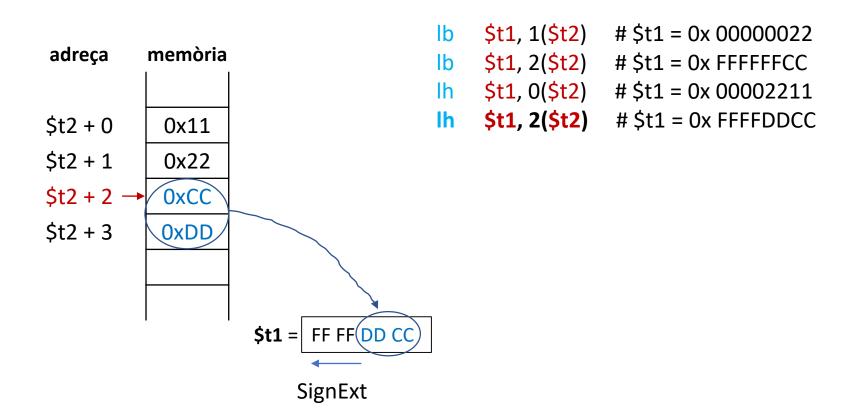
- sb rt, off16(rs) # $Mem_{byte}[rs + SignExt(off16)] \leftarrow rt_{7..0}$
 - Copia a memòria el byte de menor pes de rt

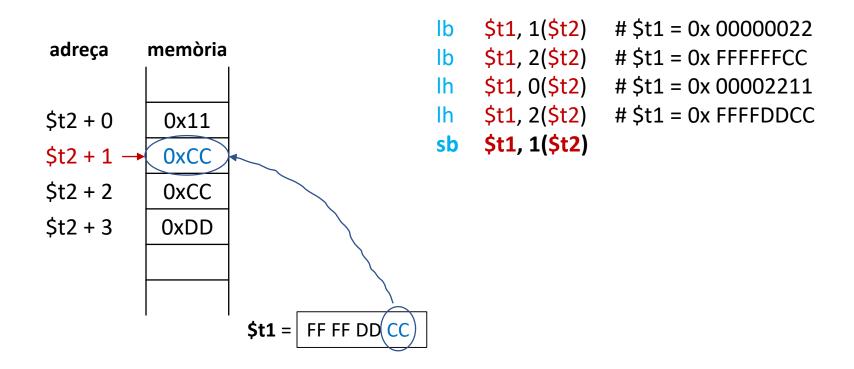












Operands en memòria halfword o byte naturals

Load Halfword Unsigned

```
lhu rt, off16(rs) # rt \leftarrow ZeroExt[Mem<sub>half</sub>[rs + SignExt(off16)]]
```

- Copia un halfword de memòria als 2 bytes de menor pes de rt
- El converteix a 4 bytes extenent zeros
- Load Byte Unsigned

```
lbu rt, off16(rs) # rt ← ZeroExt[Mem<sub>byte</sub>[rs + SignExt(off16)]]
```

- Copia un byte de memòria al byte de menor pes de rt
- El converteix a 4 bytes extenent zeros

Operands en memòria halfword o byte naturals

Load Halfword Unsigned

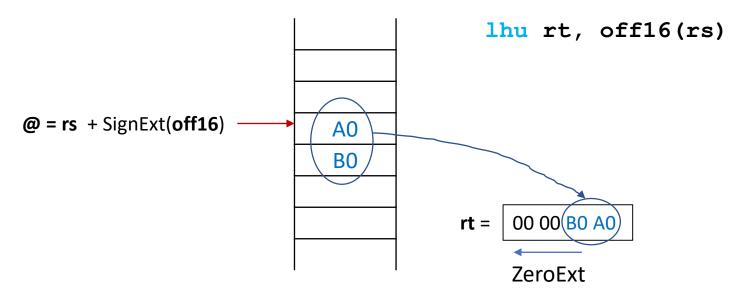
```
lhu rt, off16(rs) # rt 		 ZeroExt[Mem<sub>half</sub>[rs + SignExt(off16)]]
```

- Copia un halfword de memòria als 2 bytes de menor pes de rt
- El converteix a 4 bytes extenent zeros

Load Byte Unsigned

```
lbu rt, off16(rs) # rt ← ZeroExt[Mem<sub>byte</sub>[rs + SignExt(off16)]]
```

- Copia un byte de memòria al byte de menor pes de rt
- El converteix a 4 bytes extenent zeros



Operands en memòria doubleword

- Com s'accedeix a una paraula de 64 bits (long long)?
 - Declaració en C:

```
long long x = 0x7766554433221100;
```

En assemblador:

```
.data
x: .dword 0x7766554433221100
```

Codi assemblador per carregar x en \$t0-\$t1:

```
.text
...
# Suposem que $t2 conté l'adreça de memòria de x
lw $t0, 0($t2)
lw $t1, 4($t2)
```

Restricció d'alineació

- L' adreça efectiva de lw i sw ha de ser múltiple de 4
- L' adreça efectiva de lh, lhu i sh ha de ser múltiple de 2
- En cas contrari
 - Es produeix una excepció per adreça no-alineada
 - El programa acaba

Exemple:

```
.data
x: .word 0xDDCCBBAA
    .text
    ...
# Suposem que $t2 conté l'adreça de memòria de x
lw $t0, 1($t2)
```

o Adreça no-alineada! → excepció

- Simplifiquen operacions frequents i faciliten l'escriptura i depuració del codi
- L'assemblador l'expandeix a una o vàries instruccions

```
.data
x: .word 0
x: .word 0
y: .word 69, 70

.text
move $t1, $t2

.data
x: .word 0
x: .word 0
y: .word 69, 70
```

- Simplifiquen operacions frequents i faciliten l'escriptura i depuració del codi
- L'assemblador l'expandeix a una o vàries instruccions

```
.data
x: .word 0
y: .word 69, 70

.text
move $t1, $t2

it $t1, 100

.data
x: .word 0
y: .word 69, 70

.text
addu $t1, $t2, $zero

addiu $t1, $zero, 100
```

- Simplifiquen operacions frequents i faciliten l'escriptura i depuració del codi
- L'assemblador l'expandeix a una o vàries instruccions

```
.data
x: .word 0
y: .word 69, 70

.text
move $\frac{1}{2}$$ $\frac
```

- Simplifiquen operacions frequents i faciliten l'escriptura i depuració del codi
- L'assemblador l'expandeix a una o vàries instruccions

```
.data
                                         .data
x: .word 0
                                      x: .word 0
                                      y: .word 69, 70
y: .word 69, 70
                                         .text
   .text
move $t1, $t2
                                      addu $t1, $t2, $zero
li $t1, 100
                                      addiu $t1, $zero, 100
Ιi
  $t1, 0x075080AA
                                      lui $at, 0x0750
                                      ori
                                             $t1, $at, 0x80AA
       $t1, y # $t1\leftarrow0x10010004 \rightarrow lui $at, 0x1001
la
                                      ori
                                             $t1, $at, 0x0004
```

- Simplifiquen operacions frequents i faciliten l'escriptura i depuració del codi
- L'assemblador l'expandeix a una o vàries instruccions

```
.data
                               .data
x: .word 0
                            x: .word 0
                            y: .word 69, 70
y: .word 69, 70
                               .text
  .text
move $t1, $t2
                            addu $t1, $t2, $zero
li $t1, 100
                            addiu $t1, $zero, 100
lί
  $t1, 0x075080AA
                            lui $at, 0x0750
                            ori
                                  $t1, $at, 0x80AA
     lui
                                 $at, 0x1001
la 
                            ori
                                  $t1, $at, 0x0004
     la
                                  $t1, $at, 0x0008
                            ori
```

Exercici

o Tradueix a MIPS el programa en C

Exercici

Tradueix a MIPS el programa en C

```
short v[3] = \{-1, 32, 15\}, sum; // v, sum són globals
 void main() {
          sum = v[0] + v[1] + v[2] + 91;

    Traducció

.data
v: .half -1, 32, 15
sum:.half 0
  .text
main:
   la $t0, v
   lh $t1, 0($t0) # v[0]
   1h $t2, 2($t0) # v[1]
   addu $t1, $t1, $t2
   lh $t3, 4($t0) # v[2]
   addu $t1, $t1, $t3
   addiu $t1, $t1, 91
   la $t0, sum
   sh $t1, 0($t0)
```

Sistemes de representació binària

- Naturals i Enters (repàs)
- Caràcters
- Instruccions MIPS

Denotem:

```
    x<sub>u</sub> = valor natural a representar
    X = X<sub>n-1</sub> ... X<sub>1</sub> X<sub>0</sub> = vector de n bits que representa x<sub>u</sub>
```

Trobar quin número natural x_u està representat per X

$$x_u = \sum_{i=0}^{n-1} X_i \cdot 2^i$$

• Exemple: per a X = 00011001_2 $x_{11} = 1.2^4 + 1.2^3 + 1.2^0 = 25$

Denotem:

- x_u = valor natural a representar
 X = X_{n-1} ... X₁ X₀ = vector de n bits que representa x_u
- Trobar quin número natural x_u està representat per X

$$x_u = \sum_{i=0}^{n-1} X_i \cdot 2^i$$

- Exemple: per a X = 00011001_2 $x_{11} = 1 \cdot 2^4 + 1 \cdot 2^3 + 1 \cdot 2^0 = 25$
- Rang de representació: x_u ∈ [0, 2ⁿ-1]
 - ∘ Per exemple, per a n=8 és: x_{ij} ∈ [0, 255]
- Extensió del rang: afegir zeros a l'esquerra
 - $_{\circ}$ P.ex. extendre X=1001₂ a 8 bits \rightarrow X=00001001₂

Trobar la representació X per al natural x_u

$$X_0 = x_u \mod 2$$

 $X_1 = Q_0 \mod 2$

$$X_{n-2} = Q_{n-3} \mod 2$$

$$X_{n-1} = Q_{n-2} \mod 2$$

$$Q_0 = x_u \text{ div } 2$$

$$Q_1 = Q_0 \text{ div } 2$$

$$Q_{n-2} = Q_{n-3} \text{ div } 2$$

Trobar la representació X per al natural x_u

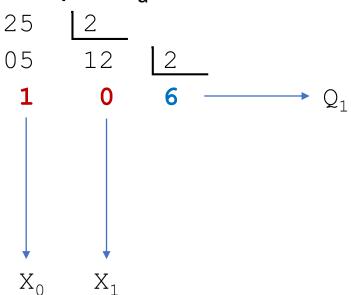
$$X_0 = x_u \mod 2$$
 $Q_0 = x_u \operatorname{div} 2$ $Q_1 = Q_0 \operatorname{div} 2$... $Q_{n-2} = Q_{n-3} \operatorname{div} 2$ $Q_{n-2} = Q_{n-3} \operatorname{div} 2$ $Q_{n-2} = Q_{n-3} \operatorname{div} 2$

• Exemple: $x_{ij} = 25$

Trobar la representació X per al natural x_u

$$X_0 = x_u \mod 2$$
 $Q_0 = x_u \operatorname{div} 2$ $Q_1 = Q_0 \operatorname{div} 2$... $Q_{n-2} = Q_{n-3} \operatorname{div} 2$ $Q_{n-2} = Q_{n-3} \operatorname{div} 2$ $Q_{n-2} = Q_{n-3} \operatorname{div} 2$

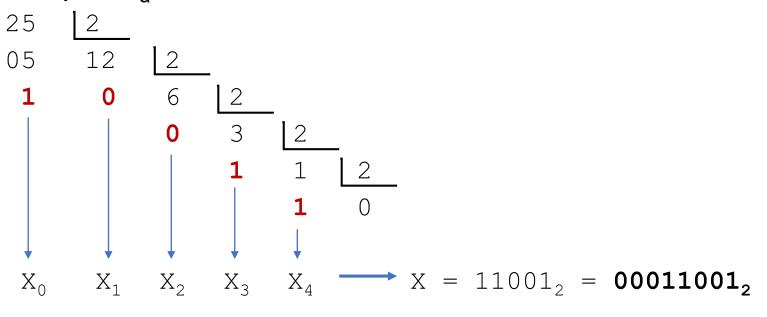
• Exemple: $x_{ij} = 25$



Trobar la representació X per al natural x_u

$$X_0 = x_u \mod 2$$
 $Q_0 = x_u \operatorname{div} 2$ $Q_1 = Q_0 \operatorname{div} 2$... $Q_{n-2} = Q_{n-3} \operatorname{div} 2$ $Q_{n-2} = Q_{n-3} \operatorname{div} 2$ $Q_{n-2} = Q_{n-3} \operatorname{div} 2$

• Exemple: $x_{ij} = 25$



Variables de tipus natural (unsigned integers) en C

En Assemblador MIPS

```
.data
var1:    .byte 0
var2:    .half 0
var3:    .word 0
var4:    .dword 0
```

• Denotem:

- x_s = valor enter a representar
 X = X_{n-1} ... X₁ X₀ = vector de n bits que representa x_s
 x₁ = natural representat per X (valor explícit de x_s)
- Trobar quin número enter x_s està representat per X

$$x_S = -X_{n-1} \cdot 2^{n-1} + \sum_{i=0}^{n-2} X_i \cdot 2^i$$

○ O bé: interpretar-lo com a natural x_u, i restar-li 2ⁿ si X_{n-1}=1

$$x_s = x_n - X_{n-1} \cdot 2^n$$

Denotem:

- x_s = valor enter a representar
 X = X_{n-1} ... X₁ X₀ = vector de n bits que representa x_s
 x₁₁ = natural representat per X (valor explícit de x_s)
- Trobar quin número enter x_s està representat per X

$$x_S = -X_{n-1} \cdot 2^{n-1} + \sum_{i=0}^{n-2} X_i \cdot 2^i$$

○ O bé: interpretar-lo com a natural x_u, i restar-li 2ⁿ si X_{n-1}=1

$$x_s = x_u - X_{n-1} \cdot 2^n$$

• Exemple: per a $X = 10011001_2$

$$x_s = -1.2^7 + 1.2^4 + 1.2^3 + 1.2^0 = -128 + 25 = -103$$

○ O bé, interpretar-lo com a natural x₁₁, i restar-li 2⁸

$$x_u = 1 \cdot 2^7 + 1 \cdot 2^4 + 1 \cdot 2^3 + 1 \cdot 2^0 = 153$$

 $x_s = x_u - 1 \cdot 2^8 = 153 - 256 = -103$

- Trobar la representació X per a l'enter x_s (en 2 passos)
 - 1. Trobar el valor explícit: $x_u = \begin{cases} x_s, & \text{si } x_s \ge 0 \\ x_s + 2^n, & \text{altrament} \end{cases}$
 - 2. Representar x_u com a natural (divisions successives)
- Exemple: representar amb 8 bits $x_s = -124$
 - 1. $x_u = x_s + 2^8 = -124 + 256 = 132$
 - 2. Divisions succ. \rightarrow el natural 132 es representa X = 10000100₂

- Trobar la representació X per a l'enter x_s (en 2 passos)
 - 1. Trobar el valor explícit: $x_u = \begin{cases} x_s, & \text{si } x_s \ge 0 \\ x_s + 2^n, & \text{altrament} \end{cases}$
 - 2. Representar x_u com a natural (divisions successives)
- Exemple: representar amb 8 bits $x_s = -124$
 - 1. $x_u = x_s + 2^8 = -124 + 256 = 132$
 - 2. Divisions succ. \rightarrow el natural 132 es representa X = 10000100₂
- Rang de representació: x_s ∈ [-2ⁿ⁻¹, 2ⁿ⁻¹ 1]
 - Per exemple, per a n=8 és: $x_s \in [-128, 127]$
- Extensió del rang: replicar el bit de signe X_{n-1} a l'esquerra
 - $_{\circ}$ Per exemple, extendre X=1001₂ a 8 bits: X=11111001₂

Regla del canvi de signe: Complementar bits i sumar 1

```
• Exemple: canviar signe de X = 01100011_2 (és l'enter x_s = 99)

X' = \overline{X} + 1

= 100111100_2 + 1

= 100111101_2 (representa l'enter x'_s = -99)
```

• Variables de tipus enter en Ca2 (signed integers) en C

En Assemblador MIPS (amb signe o sense, no canvia)

```
var1:    .byte 0
var2:    .half 0
var3:    .word 0
var4:    .dword 0
```

- Trobar quin número enter x_s està representat per X
 - ∘ Interpretar X com a natural x_u , i restar-li 2^n –1 si X_{n-1} = 1

$$x_s = x_u - X_{n-1} \cdot (2^n - 1)$$

- Trobar la representació X per a l'enter x_s (en 2 passos)
 - 1. Trobar el valor explícit: $x_u = \begin{cases} x_s, & si \ x_s \ge 0 \\ x_s + (2^n 1), & si \ x_s \le 0 \end{cases}$
 - 2. Representar x_u com a natural (divisions successives)

- Trobar quin número enter x_s està representat per X
 - Interpretar X com a natural x_u , i restar-li 2^n –1 si X_{n-1} = 1

$$x_s = x_u - X_{n-1} \cdot (2^n - 1)$$

- Trobar la representació X per a l'enter x_s (en 2 passos)
 - 1. Trobar el valor explícit: $x_u = \begin{cases} x_s, & si \ x_s \ge 0 \\ x_s + (2^n 1), & si \ x_s \le 0 \end{cases}$
 - 2. Representar x_u com a natural (divisions successives)
- Rang de representació simètric: x_s ∈ [-(2ⁿ⁻¹-1), 2ⁿ⁻¹-1]
 - Per exemple, per a n=8 és: $x_s \in [-127, +127]$
 - $_{\circ}$ 2 representacions per al zero! $00000000_2 = 111111111_2$
- Canvi de signe: complementar bits
 - Exemple: canviar de signe $X = 01100011_2$ (és l'enter $x_s = 99$) $X' = \overline{X} = 10011100_2$ (representa l'enter $x'_s = -99$)

Enters en Signe i Magnitud (repàs)

signe	e magnitud

- Trobar quin número enter x_s està representat per X
 - El bit de major pes indica el signe (0 positiu, 1 negatiu)
 - Els n-1 bits restants codifiquen el valor absolut (natural)
- Trobar la representació X per a l'enter x_s
 - Codificar el signe en el bit de major pes
 - Codificar el valor absolut (natural) en els n-1 bits restants
 - $_{\circ}$ 2 representacions per al zero! $00000000_2 = 100000000_2$

Enters en Signe i Magnitud (repàs)

signe	magnitud

- Trobar quin número enter x_s està representat per X
 - El bit de major pes indica el signe (0 positiu, 1 negatiu)
 - Els n-1 bits restants codifiquen el valor absolut (natural)
- Trobar la representació X per a l'enter x_s
 - Codificar el signe en el bit de major pes
 - Codificar el valor absolut (natural) en els n-1 bits restants
 - $_{\circ}$ 2 representacions per al zero! $00000000_2 = 100000000_2$
- Rang de representació simètric: $x_s \in [-(2^{n-1}-1), 2^{n-1}-1]$
- Canvi de signe: complementar el bit de major pes

Enters en excés a E

- Trobar quin número enter x_s està representat per X
 - 1. Interpretar X com a natural x_u
 - 2. Restar-li l'excés E (normalment, E = $2^{n-1} 1$)

$$x_S = x_u - (2^{n-1} - 1)$$

- Trobar la representació X per a l'enter x_s
 - 1. Trobar el valor explícit, sumant-li l'excés E $x_n = x_s + (2^{n-1} 1)$
 - 2. Representar-lo com a natural (divisions successives)

Enters en excés a E

- Trobar quin número enter x_s està representat per X
 - 1. Interpretar X com a natural x_u
 - 2. Restar-li l'excés E (normalment, E = $2^{n-1} 1$)

$$x_S = x_u - (2^{n-1} - 1)$$

- Trobar la representació X per a l'enter x_s
 - 1. Trobar el valor explícit, sumant-li l'excés E $x_n = x_s + (2^{n-1} 1)$
 - 2. Representar-lo com a natural (divisions successives)
- Rang de representació: $x_s \in [-(2^{n-1}-1), 2^{n-1}]$
 - Per exemple, per a n=8 és $x_s \in [-127, 128]$
- Propietat
 - Es poden comparar amb el mateix circuit que els naturals
 - S'usen per a l'exponent en números de coma flotant (Tema 5)

Caràcters

- Necessitat de representar símbols tipogràfics
 - Ja des dels temps de la telegrafia (s.XIX, codi Morse)
- Codificació = correspondència símbols ↔ números
 Unicode, EBCDIC, ASCII, etc.
- Codi ASCII de 7 bits (1963)
 - Els codis 0 a 31 són de control (no imprimibles)

codi	símbol	en C i MIPS
0x00	null	' \0'
0x09	TAB	'\t'
0x0A	LF	'\n'
0x0D	CR	'\r'
0x20	space	٤ ٤

codi	simbol	en C i MIPS
0x30	0	' 0'
0x31	1	'1'
0x41	A	'A'
0x42	В	'В'
0x61	a	ʻa'
0x62	ь	'b'

Caràcters

- Alguns codis i propietats a recordar
 - Dígits decimals (del '0' al '9')
 - a partir del codi 48 (0x30)
 - Majúscules ordenades (de 'A' a 'Z')
 - a partir del codi 65 (0x41)
 - Minúscules ordenades (de 'a' a 'z')
 - a partir del codi 97 (0x61)
 - Observar que de majúscules a minúscules sols canvia el bit 5
 cmin = cmaj + 32;
 - Representar com a caràcter un dígit decimal n (0 ≤ n < 10)
 cdigit = '0' + n;
- Alguns caràcters més
 - Caràcter null ('\0'), té codi 0 (0x00)
 - Caràcter espai (' '), té codi 32 (0x20)
 - Caràcter salt de línia ('\n'), té codi 10 (0x0A)

Caràcters

Variables de tipus caràcter, en C

(si es declara *unsigned* no canvia res, ja que el bit 8 val 0)

En Assemblador MIPS

```
.data
lletra: .byte 'R'
```

Programa en C

```
    Programa en MIPS
```

.data

```
cars: .byte 'A', '1'
n: .byte 7
 .text
main:
   la $t0, cars
   lb $t1, 0($t0)
   addiu $t1, $t1, 32
   sb $t1, 0($t0)
   la $t2, n
   lb $t1, 0($t2)
   addiu $t1, $t1, '0'
   sb $t1, 1($t0)
```

Format de les instruccions MIPS

• Sols 3 formats: R (registre), I (immediate), J (jumps)

	6 bits	5 bits	5 bits	5 bits	5 bits	6 bits	
R	opcode	rs	rt	rd	shamt	funct	
I	opcode	rs	rt	imm16			
J	opcode	target					

Alguns exemples

		6 bits	5 bits	5 bits	5 bits	5 bits	6 bits	
addu	rd, rs, rt	R	0x00	rs	rt	rd	0x00	0x21
sra	rd, rt, shamt	R	0x00	rs	rt	rd	shamt	0x03
addiu	rt, rs, imm16	I	0x08	rs	rt	imm16		
lui	rt, imm16	Ι	0x0F	0x00	rt	imm16		
lw	rt, offset16(rs)	I	0x23	rs	rt	offset16		
jal	target	J	0x03	target				

Exercici

• Codificar en binari la instrucció⁽¹⁾:

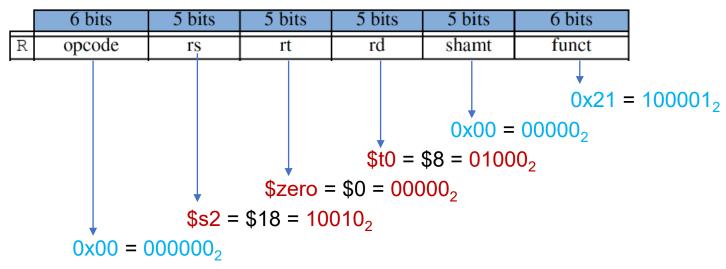
addu \$t0, \$s2, \$zero

⁽¹⁾ Consulteu sintaxi, opcodes i números de registre al document <u>Instruccions MIPS i macros MARS</u> penjat a la web

Exercici

Codificar en binari la instrucció⁽¹⁾:

- Solució
 - o Sintaxi: addu rd, rs, rt
 - o Format:



Codificació:

 $000000 \ 10010 \ 00000 \ 01000 \ 00000 \ 100001_2 = 0x02404021$

(1) Consulteu sintaxi, opcodes i números de registre al document <u>Instruccions MIPS i macros MARS</u> penjat a la web

Vectors i punters

Vectors

- Agrupació unidimensional de N elements del mateix tipus
 - Elements identificats per un índex ε [0, N-1]
 - S'emmagatzemen en posicions consecutives de memòria
 - En MIPS, han de respectar les regles d'alineació
 - Si la mida és potència de 2, alineant el primer s'alineen tots

Declaració de vectors

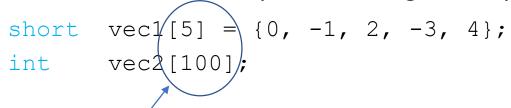
Declaracions en C (variables globals)

```
short vec1[5] = {0, -1, 2, -3, 4};
int vec2[100];
```

La dimensió: ha de ser una constant

Declaració de vectors

Declaracions en C (variables globals)



La *dimensió*: ha de ser una constant

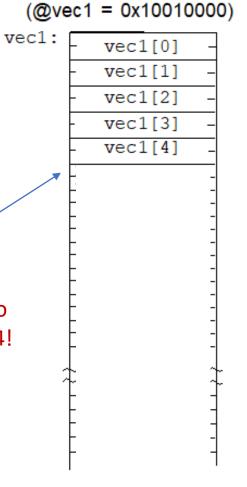
Codi en MIPS

.data

vec1: .half 0, -1, 2, -3, 4

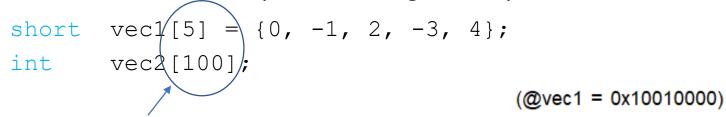
vec2: .space 400

L'adreça 0x1001000A no és múltiple de 4!



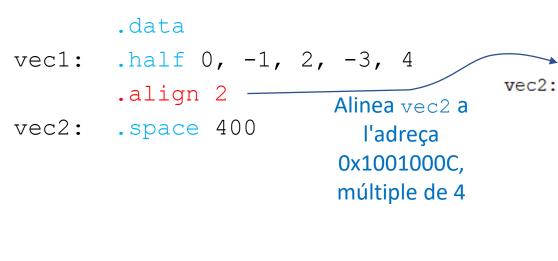
Declaració de vectors

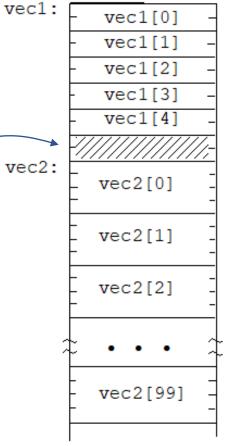
Declaracions en C (variables globals)



La *dimensió*: ha de ser una constant

Codi en MIPS



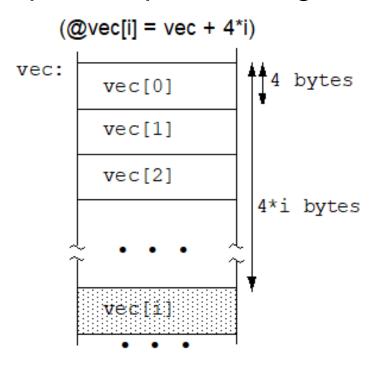


Accés a un element d'un vector

- Per accedir a l'element i-èssim
 - N'hem de calcular l'adreça
 - Suposant elements de mida = T bytes

$$@vec[i] = @vec[0] + i \cdot T$$

• Exemple (suposem que vec és global i *T*=4)



Accés a un element d'un vector

• Exemple, amb un índex constant: vec[3]

```
int vec[100];
main() {
    int x; // suposem que x es guarda en $t1
    x = vec[3];
}
```

Traducció de la sentència a MIPS

```
la $t0, vec
lw $t1, 12($t0)

obé...

la $t0, vec+12
lw $t1, 0($t0)
```

Accés a un element d'un vector

• Exemple, amb un índex variable: vec[i]

```
int vec[100];
main() {
    int x, i; // suposem x, i guardats en $t1, $t2
    x = vec[i];
}
```

Traducció de la sentència a MIPS

```
la $t0, vec # $t0 = @vec[0]
sll $t3, $t2, 2 # $t3 = i*4
addu $t0, $t0, $t3 # $t0 = @vec[0] + i*4
lw $t1, 0($t0) # x = vec[i]
```

Strings o cadenes de caràcters

- Són vectors amb un nombre variable de caràcters
- Emmagatzemament
 - En Java: com una tupla (un enter i un vector de caràcters)
 - En C: vector de caràcters amb sentinella (caràcter null = '\0')
- Declaracions en C (formes equivalents)

Traducció a MIPS (formes equivalents)

Strings o cadenes de caràcters

Exemple en C

```
char nom[80];
main() {
    int num = 0;
    ...
    while (nom[num] != '\0') num +=1;
}
```

Traducció a MIPS

Punters

- Punter: Variable que conté una adreça de memòria
 - o 32 bits en MIPS
 - Si el punter p conté l'adreça de la variable v...
 - ... diem que *p apunta a v*

Declaració de punters

Declaració dels punters p1, p2 i p3 en C:

```
int *p1, *p2;  // p1, p2 apunten a variables de tipus int
char *p3;  // p3 apunta a variables de tipus char
```

Alerta amb els tipus!

- Declaracions de p1, p2 i p3 en MIPS
 - Si són globals:

```
.data
p1: .word 0
p2: .word 0
p3: .word 0
```

- Si són locals:
 - no cal reservar, sols decidir en quin registre es guardaran

Inicialització d'un punter

- Assignant-li un altre punter del mateix tipus
- Assignant-li l'adreça d'una variable (operador &, en C)
 - Si suposem que el punter és global (pglob):

```
char a = 'E';
char b = 'K';
char *pglob = &a;

.data
a: .byte 'E'
b: .byte 'K'
pglob: .word a
```

Inicialització d'un punter

- Assignant-li un altre punter del mateix tipus
- Assignant-li l'adreça d'una variable (operador &, en C)
 - Si suposem que el punter és global (pglob):

```
char a = 'E';
char b = 'K';
char *pglob = &a;

void f() {
    pglob = &b;
    ...
}
```

```
.data
a: .byte 'E'
b: .byte 'K'
pglob: .word a

.text
f:
    la $t0, b
    la $t1, pglob
    sw $t0, 0($t1)
```

Inicialització d'un punter

- Assignant-li un altre punter del mateix tipus
- Assignant-li l'adreça d'una variable (operador &, en C)
 - Si suposem que el punter és global (pglob):

Si suposem que el punter és local (ploc) i el guardem en \$t5:

Operació desreferència (indirecció) de punters

Desreferència:

- Consisteix a accedir a l'adreça de memòria apuntada pel punter
- En C, usant l'operador * precedint el punter
 - *p significa "contingut de l'adreça de memòria apuntada per p"
 - No confondre amb el símbol * usat per declarar punters!

Operació desreferència (indirecció) de punters

Desreferència:

- Consisteix a accedir a l'adreça de memòria apuntada pel punter
- En C, usant l'operador * precedint el punter
 - *p significa "contingut de l'adreça de memòria apuntada per p"
 - No confondre amb el símbol * usat per declarar punters!
- Exemple: suposem que el punter és global (pglob):

```
char a = 'E';
char *pglob = &a;

void ff() {
    char tmp; // guardat en $t0
    tmp = *pglob;
    ...
}

.data
.byte 'E'
pglob: .word a

.text

ff:
    tmp = *pglob;
    la $t1, pglob
lw $t5, 0($t1)
lb $t0, 0($t5)
```

Operació desreferència (indirecció) de punters

Desreferència:

- Consisteix a accedir a l'adreça de memòria apuntada pel punter
- En C, usant l'operador * precedint el punter
 - *p significa "contingut de l'adreça de memòria apuntada per p"
 - No confondre amb el símbol * usat per declarar punters!
- Exemple: suposem que el punter és global (pglob):

```
char a = 'E';
char *pglob = &a;

void ff() {
    char tmp; // guardat en $t0
    tmp = *pglob;
    ...
}

.data
a:    .byte 'E'
pglob: .word a

.text

ff:

tmp = *pglob;
la $t1, pglob
lw $t5, 0($t1)
lb $t0, 0($t5)
```

Exemple: suposem que el punter és local (ploc), guardat en \$t5:

```
void gg() {
    char tmp, *ploc; //en $t0,$t5
        tmp = *ploc;
}
.text
gg:
1b $t0, 0($t5)
```

Operació aritmètica de punters

Aritmètica de punters

- Suma d'un punter p més un enter N
- \circ Dóna com a resultat un altre punter q del mateix tipus: q = p + N
- q apunta a una adreça situada N elements més endavant
- \rightarrow Si els elements tenen mida T, cal sumar N^*T bytes a l'adreça

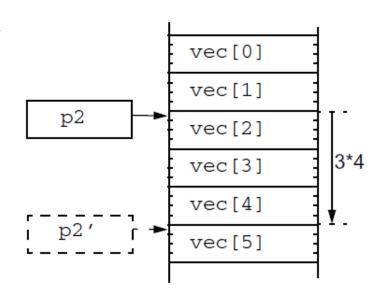
Exemple

Suposem que p2 s'ha inicialitzat així

int
$$*p2 = \&vec[2];$$

o i executem la següent operació

$$p2 = p2 + 3;$$



Operació aritmètica de punters

Exemple

Suposem les següents declaracions en C

```
char *p1;  // guardat en $t1
int *p2;  // guardat en $t2
long long *p3;  // guardat en $t3
```

I les sentències...

```
p1 = p1 + 3;
p2 = p2 + 3;
p3 = p3 + 3;
```

La traducció a MIPS seria...

```
addiu $t1, $t1, 3
addiu $t2, $t2, 12
addiu $t3, $t3, 24
```

En C, un vector és en realitat un punter que apunta al seu primer element

Per exemple

```
int vec[100];
int *p;
```

- Els tipus de p i vec són equivalents!
 - Però vec és constant, sempre apunta al mateix element
 - o En canvi, p és variable, pot apuntar a elements diferents

En C, un vector és en realitat un punter que apunta al seu primer element

- En conseqüència
 - 1. Podem escriure

```
p = vec;  // Fem que p apunti a vec[0]
```

En C, un vector és en realitat un punter que apunta al seu primer element

- En conseqüència
 - 1. Podem escriure

```
p = vec;  // Fem que p apunti a vec[0]
```

2. Als punters també els podem aplicar l'operador []

```
p[8] = 3; // assignem un 3 a vec[8]
```

En C, un vector és en realitat un punter que apunta al seu primer element

- En conseqüència
 - 1. Podem escriure

```
p = vec;  // Fem que p apunti a vec[0]
```

2. Als punters també els podem aplicar l'operador []

```
p[8] = 3; // assignem un 3 a vec[8]
```

3. Als vectors també els podem aplicar l'operador *

```
*vec = 5; // assignem un 5 a vec[0]
```

En C, un vector és en realitat un punter que apunta al seu primer element

- En conseqüència
 - 1. Podem escriure

```
p = vec;  // Fem que p apunti a vec[0]
```

2. Als punters també els podem aplicar l'operador []

```
p[8] = 3; // assignem un 3 a vec[8]
```

3. Als vectors també els podem aplicar l'operador *

```
*vec = 5; // assignem un 5 a vec[0]
```

- 4. Si p apunta a vec[0], p+i apunta a vec[i]
 - per tant, l'expressió

```
*(p+i) = 0; // assignem un 0 a vec[i]
```

- és equivalent a

```
p[i] = 0; // assignem un 0 a vec[i]
```