第4章 寻址方式

寻址方式：寻找操作数存放地址的方法，即根据指令计算操作数地址的方法。

问题：如何得到某人的住址？

- CPU如何知道操作数的地址？
- C程序中，有哪些给出地址的方式？

程序中和寻址方式相关的部分

```
main proc
    mov   eax, 0        ;  (eax)=0，用于存放累加和
    mov   ebx, 1        ;  (ebx)=1，用于指示当前的加数
lp: cmp   ebx, 100      ; 连续两条指令，等同于
                        ;  if (ebx>100) goto exit
     jg    exit
     add   eax, ebx     ;  (eax) = (eax) + (ebx)
     inc   ebx          ;  (ebx) = (ebx) +1
     jmp   lp           ;  goto lp
exit:
    invoke printf, offset lpFmt, eax
    invoke ExitProcess, 0
main  endp
end
```

.所谓寻址方式就是在指令中表达操作数的方法。

.实际上决定了常用汇编指令的语法格式。

第4章　寻址方式

一、本章的学习内容

立即寻址

寄存器寻址

直接寻址

寄存器间接寻址

变址寻址

基址加变址寻址

寻址方式的综合举例

x86机器指令编码规则

4.1 寻址方式概述

一条指令，关注的焦点有哪些？

- 执行什么操作？
- 操作数在哪里？

操作数的存放位置即存放地址
- CPU内的寄存器
- 主存
- I/O设备端口

操作数在主存时：关注段址/段选择符、段内偏移

- 操作数的类型　字节/字/双字？

寻址方式关注的要素：位置(内存/寄存器/立即数)，类型
对于存储器操作数还要关注：段选择符：偏移地址。

4.1 寻址方式概述

双操作数的指令格式　　操作符 OPD, OPS

ADD　EAX，EBX

目的操作数地址　　源操作数地址

(OPD) + (OPS) → OPD
(EAX) + (EBX) → EAX

Question：指令MOV BX，10 的寻址方式是什么？

4.2 立即寻址

- 操作数直接放在指令中，在指令的操作码后；
- 操作数是指令的一部分，位于代码段中；
- 指令中的操作数是8位、16位或32位二进制数。

使用格式： n

操作码及目的操作数寻址方式码
立即操作数 n

立即操作数只能作为源操作数。

例： MOV EAX , 12H

机器码： B8 12 00 00 00

4.2 立即寻址

- 立即操作数只能作为源操作数。
- 立即数值的大小应在另一个操作数的类型限定的取值范围内

  MOV AL, 1234H;

  error: invalid instruction operands
- 字符、字符串是立即数

  MOV AL, ’1’ MOV AL, 31H

  MOV AX, ’12’ MOV AX, 3132H
- 由常量组成的数值表达式是立即数

  MOV EAX, 3*4+5*6 MOV EAX, 42

外在的表象（表达式）-» 编译-» 内部的本质（立即数）

4.2 立即寻址

- 取一个数据段中偏移地址得到的是立即数

```
.data
    x    db  10, 20, 30, 40, 50, 60

MOV  EAX, OFFSET  X
MOV  EAX, OFFSET X + 2
```

编译器先计算 OFFSET X 得到一个立即数，
再将该立即数与 2 相加，得到的结果
相当于 x 中 数30所在单元的地址
立即数的数据类型?

4.3 寄存器寻址

使用格式： R

功　　能： 寄存器R中的内容即为操作数。

说　　明： 除个别指令外，R可为任意寄存器。

例1：DEC　BL

执行前 ： (BL)=43H

执　行： (BL) –1 = 43 H –1 = 42H → BL

执行后： (BL)=42H

Question： 操作数在哪？操作数类型是什么？

4.3 寄存器寻址

Question： 指令 MOV AX, BH 正确吗？为什么？

讨论指令: ADD EAX, BX 是否正确？为什么？

```
MOV CS, AX
MOV BX, EIP
MOV EAX, EBX+3
```

4.4 直接寻址

- 操作数在内存中；
- 操作数的偏移地址EA紧跟在指令操作码后面。

格式：变量 ；等同 [变量]

变量 + 常量 ；等同 变量[常量]

；等同 [变量 +常量]

要点：编译器对 变量+常量 进行编译时，
取变量的地址，将该地址与常量运算，
得到一个新地址存放在机器码中。

Q：操作数的类型是什么？
段地址和偏移地址怎样得到？

4.4 直接寻址


```
X    DD   10,   20,   30,   40
MOV    EAX,   X          ;   (EAX)=10
MOV    EAX,   [X]        ;   (EAX)=10
MOV    EAX,   X+4        ;   (EAX)=20
MOV    EAX,  X[4]        ;   (EAX)=20
MOV    EAX,   [X+8]      ;   (EAX)=30
MOV    EAX,   X+1        ;   (EAX)= 14000000H
```

Q：直接寻址与C语言中一维数组的访问有何异同点？

Q：如何将 x中第 0 个双字数据 +2-> EAX?

```
MOV   EAX,   X
ADD   EAX,   2
```

4.4 直接寻址

直接寻址另一种写法

格式：段寄存器名：[n]

功能：操作码的下一个字（或双字）单元
的内容为操作数的偏移地址EA。

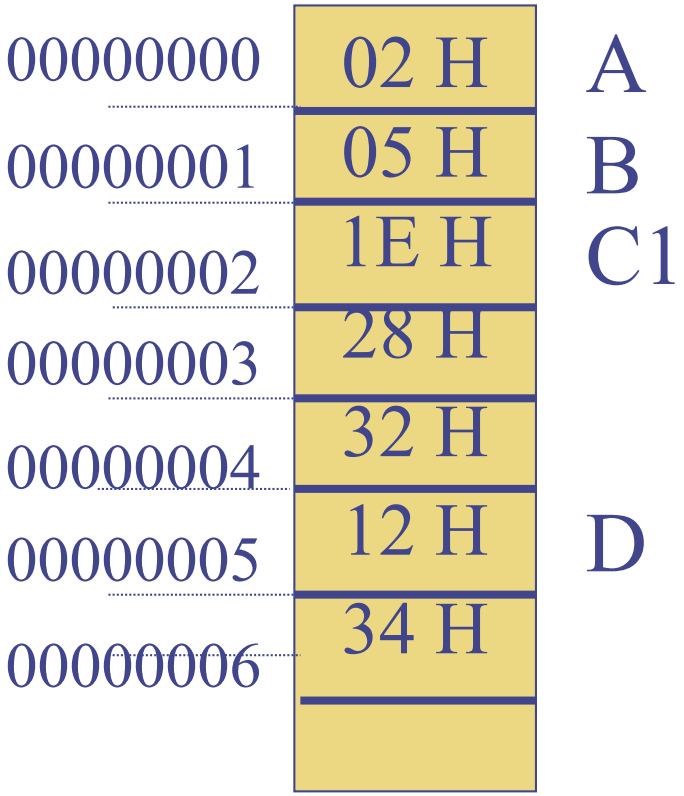
- 操作数所在的段：由段寄存器名指示
- 操作数的类型：未知

变量 或者 变量[常量] 编译后对应的值即为 n.

4.4 直接寻址


```
.DATA
A    DB   2
B    DB   5
C1   DB   30, 40, 50
D    DW   3412H
```

C 不能作为变量名



(1) MOV AH，B
(2) MOV CX, D
(3) MOV AL, C1+1
(4) MOV AX, WORD PTR A
(5) MOV EAX,DWORD PTR A

(AH)=5
(CX)=3412H
(AL)=40
(AX)=0502H
(EAX)=281E0502H

4.5 寄存器间接寻址

格式：[R]

功能：操作数在内存中，操作数的偏移地址在寄存器R中。即（R)为操作数的偏移地址.

- R 可以是：
  8个32位通用寄存器中的任意一个
  EAX, EBX, ECX, EDX, ESI, EDI, ESP, EBP

- 操作数的偏移地址在指令指明的寄存器中
- 操作数所在的段是？
  扁平内存管理模式下，（DS）=（SS）
- 操作数的类型：未知，如何确定？

4.5 寄存器间接寻址

例1：MOV AX, [ESI]

执行前 (AX)=0005H

(ESI) =00000020H

DS:(00000020H)=1234H

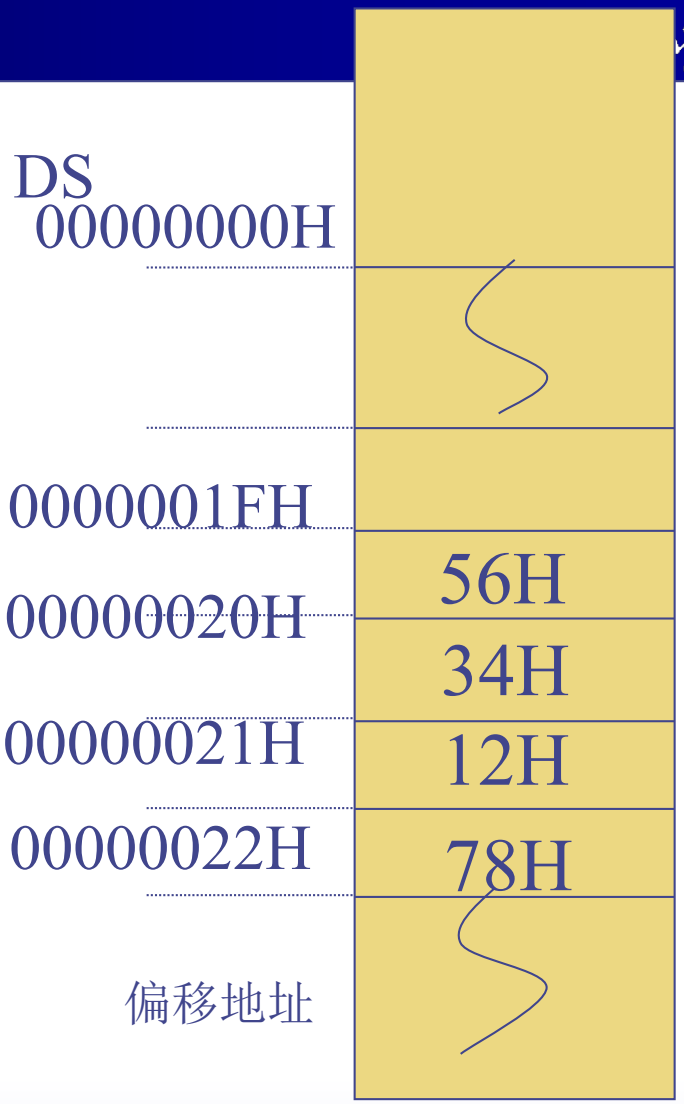
执行后 (AX)= 1234H

(ESI) = 00000020H

问：MOV CL, [ESI]

(CL) = ?

操作数的类型是如何确定的?



4.5 寄存器间接寻址

例2：MOV AH, [EBP]

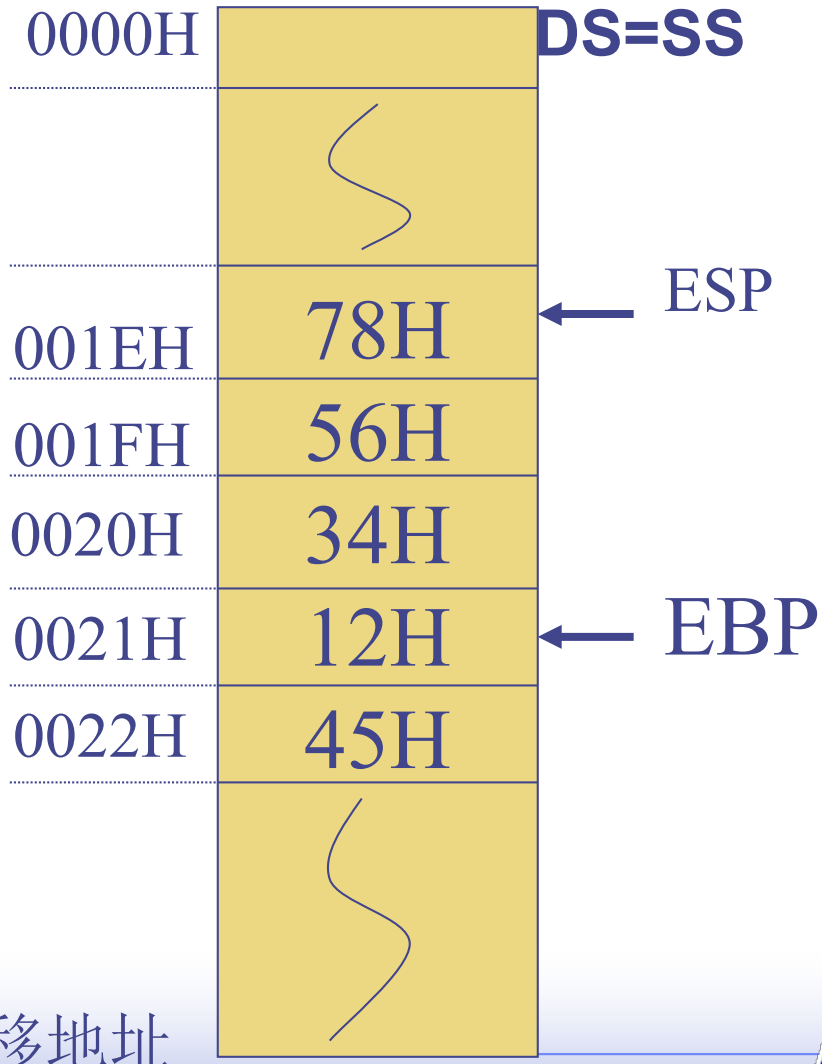
执行前 (AX)=0005H

(EBP) =21H

SS:(EBP) =12H

执行后 (AX) =1205H

(EBP) =21H



4.5 寄存器间接寻址

比较：

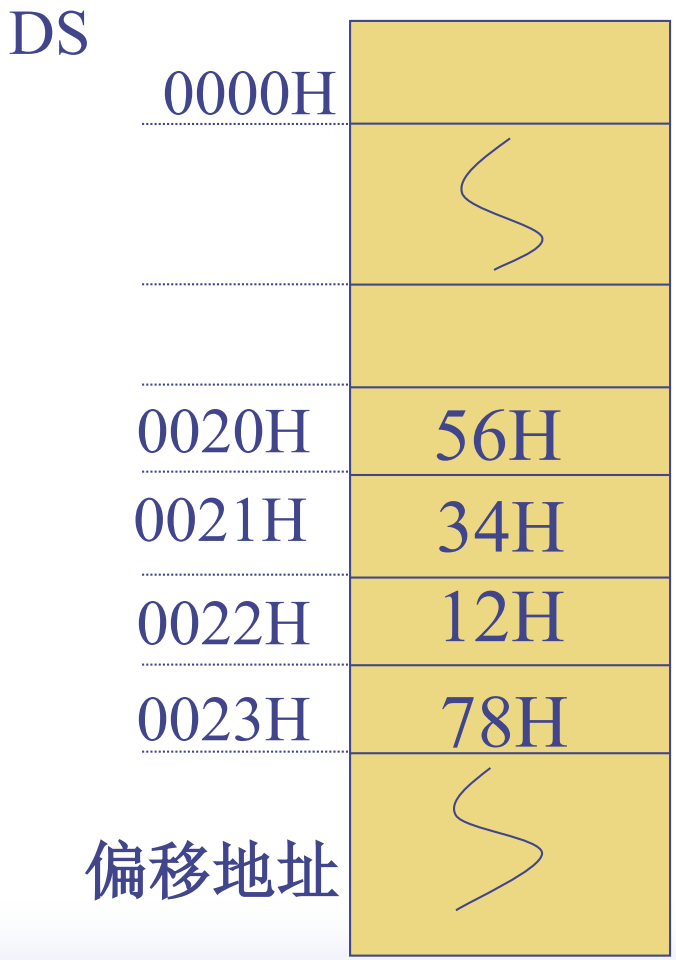
```
MOV    EAX，EBX
MOV    EAX, [EBX]
```

EAX	78123456H
EBX	00000020H

比较：

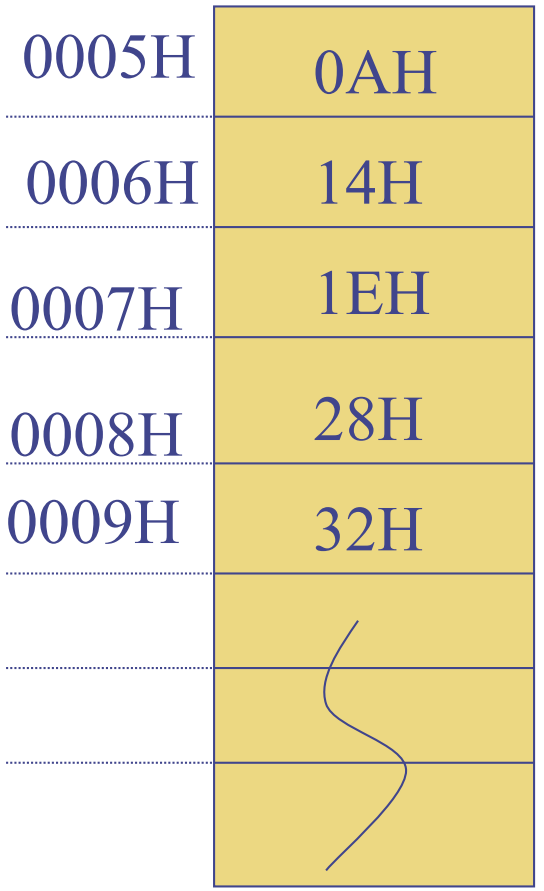
~~MOV AX，EBX~~

```
MOV   AX, [EBX]
```



4.5 寄存器间接寻址

例：设 BUF DB 10,20,30,40,50
即以BUF为首址的字节区中存放有5个数据，求它们的和。



算法分析：……

和？ AL
循环次数？ ECX
数据位置？ EBX

EBX 0005 ([EBX])

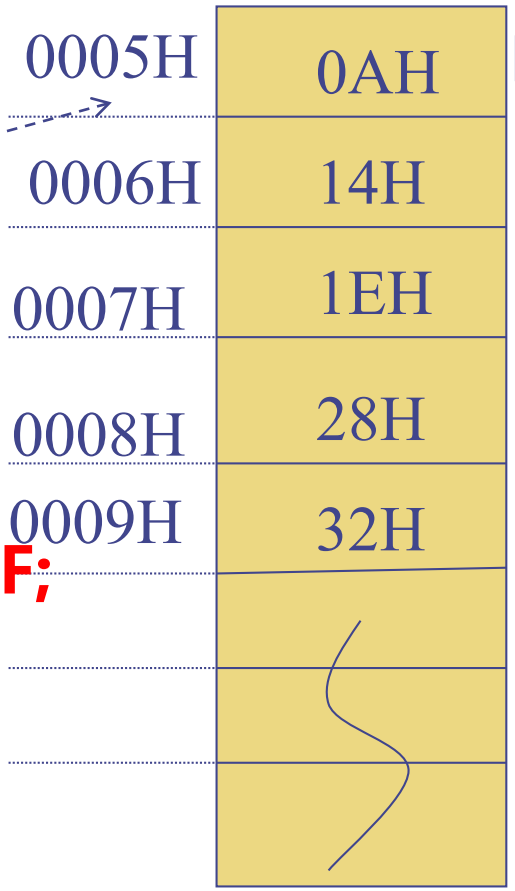
共同特点：单元中的内容无规律，
但单元之间的地址有规律。

4.5 寄存器间接寻址


```
char  buf[5]={10,20,30,40,50} ;

 char  *p;

  0005
```

EBX

```
 p=buf;        MOV EBX, OFFSET BUF;

 AL=AL+*p;     ADD AL, [EBX];
 p=p+1;        ADD EBX, 1;
```



C 用指针实现存储单元内容的间接访问；

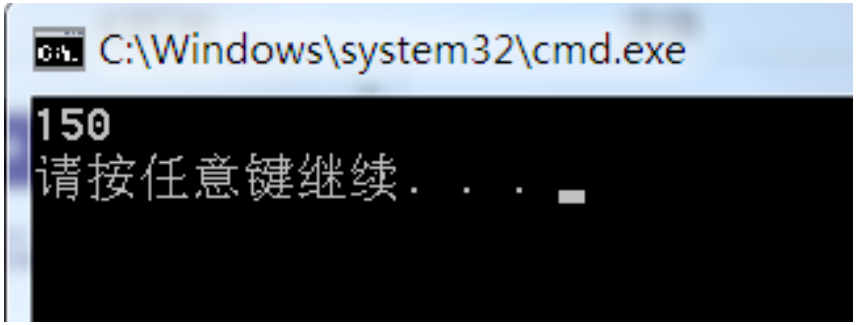
本质：寄存器间接寻址

4.5 寄存器间接寻址


```
; 程序功能：求一个数组缓冲区中 5个字节数据的和，输出和
.686
.model flat, stdcall

; ExitProcess 在 kernel32.lib中实现，原型定义如下
  ExitProcess PROTO STDCALL :DWORD
  includelib  kernel32.lib

 ; printf 在 msvcrt.lib中实现，原型定义如下
  includelib      msvcrt.lib
  printf          PROTO C :ptr sbyte, :VARARG

.DATA
  lpFmtdb"%d",0ah, 0dh, 0
  buf   db  10,20,30,40,50

.STACK 200
```

4.5 寄存器间接寻址


```
.CODE
 main    proc
    mov   eax, 0
    mov   ebx, offset buf
    mov   ecx, 0
lp: cmp   ecx, 5
    jge    exit
    add   al, [ebx]
    add   ebx, 1
    inc   ecx
    jmp   lp
exit:
    invoke printf,offset lpFmt,eax
    invoke ExitProcess, 0
main  endp
end
```



Question:

```
 add   al, [ebx]
```

能否改为：

```
 add al, ebx
```

不能，语法错误

操作数类型不匹配

Question:

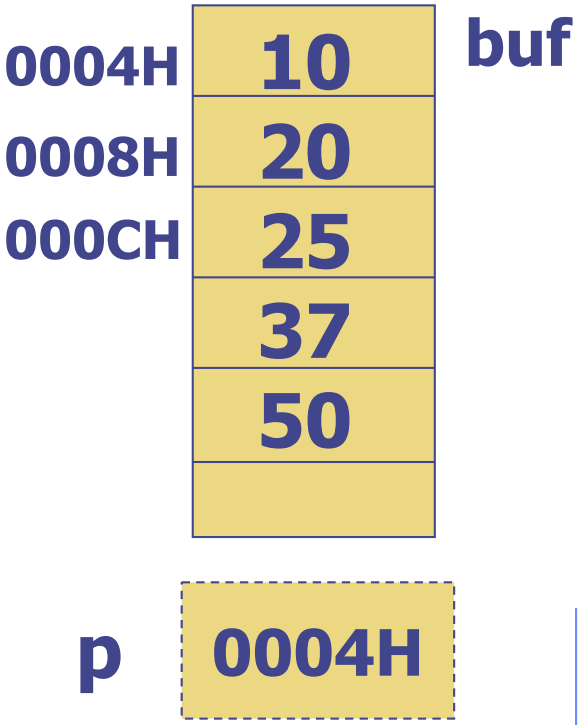
```
 add   al, [ebx]
```

能否改为：

```
 add eax, [ebx]
```

4.5 寄存器间接寻址

```
int   buf[5]={10,20,25,37,50};
int   i;
int  *p;
int   result=0;
p=buf;
for  (i=0;i<5;i++)
{
    result+=*p;
    p=p+1;
}
```



P与EBX对应：

```
MOV   EBX,   OFFSET buf
ADD   EAX,   [EBX]
ADD   EBX,    4
```

4.5 寄存器间接寻址——16位段程序

8086中分段管理模式——16位段程序:

四个寄存器：BX, BP, SI, DI

386CPU的分段管理模式——16位/32位段程序:

EAX, EBX, ECX, EDX, ESI, EDI, EBP, ESP
——任何一个或者

BX, BP, SI, DI——任何一个

4.6 变址寻址

格式：V[R×F] 或[R×F+V] 或 [R×F] +V

功能：R中的内容×F + V 为操作数的偏移地址。

- R 可以是：

  8个32位通用寄存器

  EAX, EBX, ECX, EDX, ESI, EDI, ESP, EBP

- F 可为 1，2，4，8
- V 可为 常数，变量

操作数：在存储器，段？ 偏移地址？数据类型？

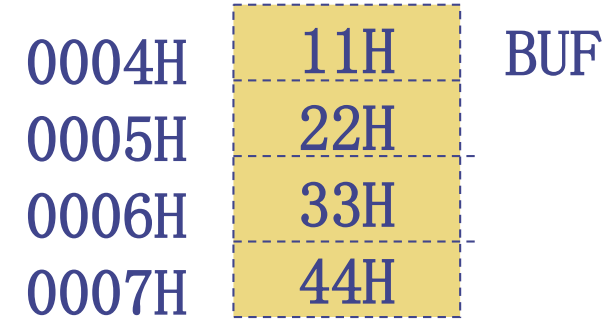
4.6 变址寻址

格式：[R×F+V]　　或　　[R×F]+V

- V 可为数值常量，也可以为一个变量。
- 当V 为变量时，取该变量对应单元的有效地址参与运算。
- **操作数所在的段 (DS) = (SS)**

例：BUF DW 2211H, 4433H

MOV AX, BUF[EBX*2]



Question:

若 （EBX）=0，　（AX）=?　2211H

若 （EBX）=1，　（AX）=?　4433H

4.6 变址寻址

- 操作数所在的段

在 VS 2109 编程中，

直接在 .DATA, .STACK, .CODE 中定义的变量是全局变量。

在 编译链接后， DS与 SS的值相同。

在 CODE 中的变量，也会变成 DS：[...]

在 子程序中，可以定义局部变量，其空间分配方式完全不同。

4.6 变址寻址

用变址寻址方式写一段程序，求BUF中双字类型数据的和

```
.686
.model flat, stdcall
 ExitProcess PROTO STDCALL :DWORD
 includelib  kernel32.lib
 includelib       msvcrt.lib
 printf PROTO C :ptr sbyte, :VARARG

.DATA
  lpFmtdb"%d",0ah, 0dh, 0
  buf    dd   10,20,30,40,50

.STACK 200
```

```
.CODE
main  proc
     MOV   EAX, 0
     MOV   EBX, 0
LP:  CMP   EBX, 5
    JGE    EXIT
    ADD    EAX, buf[EBX*4]
    ADD    EBX, 1
    JMP    LP
EXIT:
 invoke printf,offset lpFmt,EAX
    invoke ExitProcess, 0
main  endp
END
```

4.6 变址寻址


```
int   buf[5]={10,20,25,37,50} ;
int   i;
int   result=0;
for  (i=0;i<5;i++)
      result+=buf[i];
```

思考:

(EBX) 中的值就是 变量i的值，

为何C语言中的访问方式是 buf[i]，

而汇编语言中是 buf[EBX*4] ？

4.6 变址寻址

变址寻址和寄存器间接寻址有何异同？

```
        MOV   EBX , 0
        MOV   EAX , 0
LP:     CMP   EBX, 5
        JGE   EXIT
        ADD   EAX,BUF [EBX*4]
        INC    EBX
        JMP    LP
```

```
        MOV   ECX , 0
        MOV   EAX , 0
        MOV   EBX , OFFSET  BUF
LP:     CMP   ECX,5
        JGE    EXIT
        ADD   EAX, [EBX]
        ADD   EBX,4
        INC    ECX
        JMP    LP
```

4.6 变址寻址——16位段程序

8086中分段管理模式——16位段程序: V[R]

四个寄存器：BX, BP, SI, DI

V是变量时，使用的段寄存器取决于该变量定义所在的段和哪一个段寄存器建立了关联

386CPU的分段管理模式——16位/32位段程序：V[R*F]

EAX, EBX, ECX, EDX, ESI, EDI, EBP, ESP——任何一个或者

BX, BP, SI, DI(F=1)——任何一个

4.7 基址加变址寻址

格式：　[BR＋IR×F+V]

或 V[BR][IR×F]　或　V [IR×F][BR]

或 V[BR+IR×F]

功 能：操作数的偏移地址 ＝变址寄存器IR中的内容×比例因子F +位移量V+基址寄存器BR中的内容 。

EA = (IR)*F + V + (BR)

例如：MOV　EAX, -6[EDI*2][EBP]

4.7 基址加变址寻址

◆F 可为 1，2，4，8

◆当使用32位寄存器时

BR可以是 EAX，EBX，ECX，EDX，ESI，EDI，ESP，EBP 之一；

IR 可以是除ESP外的任一32位寄存器；

未带比例因子的寄存器是 BR；

当没有比例因子时，写在前面的寄存器是BR.

4.7 基址加变址寻址

特别说明：

当V中存在全局变量或标号时，用的段都是 DS。

在 VS2019中， (DS)=(SS)。

CS中的全局变量等同 DS段中的变量。

- 操作数的类型：

  若V为变量，则操作数类型为变量的类型；

  若V为常量，类型未知。

4.7 基址加变址寻址——16位段程序

8086中分段管理模式——16位段程序: V[BR+IR]

基址寄存器：BX, BP　　变址寄存器：SI，DI

V是变量时，使用的段寄存器取决于该变量定义所在的段和哪一个段寄存器建立了关联

386CPU的分段管理模式——16位/32位段程序：V[BR+IR*F]

基址寄存器：EAX, EBX, ECX，EDX, ESI, EDI, EBP, ESP
变址寄存器：除ESP外的7个32位通用寄存器

基址寄存器：BX, BP　　变址寄存器：SI，DI　F=1
V是变量时，使用的段寄存器取决于该变量定义所在的段和哪一个段寄存器建立了关联

4.8 寻址方式综合举例

寻址方式有6种。
根据操作数的存放位置，寻址方式归为3类：

立即方式
寄存器方式
存储器方式
- 寄存器间接寻址
- 变址寻址
- 基址加变址寻址
- 直接寻址

4.8 寻址方式综合举例

1. 双操作数寻址方式的规定

一条指令的源操作数和目的操作数不能同时用存储器方式。

```
MOV   AX,  BX          MOV   BUF, 0
MOV   BUF, BX          MOV   BX,   0
MOV   BX,   BUF
```

~~MOV BUF, MSG~~

~~MOV BUF, [EBX]~~

~~MOV BUF, 5[EBX]~~

~~MOV BUF, 5 [EAX + EDI *4]~~

4.8 寻址方式综合举例

2. 操作数的类型

- 寄存器寻址方式中，操作数的类型由谁决定？

  寄存器

- 立即数类型？

  不明确

- 不含变量的存储器方式所表示的操作数类型？

  未知

- 含变量的寻址方式对应的操作数类型 ？

  变量的类型

4.8 寻址方式综合举例

3. 双操作数的类型规定

- 双操作数中至少应有一个的类型是明确的；
- 若两个操作数的类型都明确，则两个的类型应相同。

(1) MOV BX, AX

(2) MOV BX, AL ✗ 两个操作数的类型不匹配

(3) MOV [EBX], 0 ✗ 两个操作数的类型不明确

属性定义算符 PTR

```
MOV   BYTE  PTR [EBX],   0
MOV   WORD PTR [EBX],   0
MOV   DWORD PTR[EBX], 0
```

4.9 x86机器指令编码规则

指令前缀	操作码	内存/寄存器操作数	索引寻址描述	地址偏移量	立即数

① 指令前缀(prefix，非必需，0个或多个字节)

② 操作码(opcode，必须，1字节～3字节)

③ 内存/寄存器操作数(ModR/M，非必需)

指明寻址方式

④ 索引寻址描述 (SIB，非必需)

指明：基址寄存器、变址寄存器、比例因子

⑤ 地址偏移量(Displacement，非必需)

⑥ 立即数(Immediate，非必需)

4.9 x86机器指令编码规则

华中科技大学

指令前缀

种类	名称	二进制码	说明
LOCK	LOCK	F0H	让指令在执行时候先禁用数据线的复用特性，用在多核的处理器上，一般很少需要手动指定
REP	REPNE/REPNZ	F2H	用 CX(16 位下)或 ECX(32 位下)或 RCX(64 位下)作为指令是否重复执行的依据
	REP/REPE/REPZ	F3H	同上
Segment Override	CS	2EH	段重载(默认数据使用 DS 段)
	SS	36H	同上
	DS	3EH	同上
	ES	26H	同上
	FS	64H	同上
	GS	65H	同上
REX	64 位	40H-4FH	x86-64 位的指令前缀，见第 18 中的介绍
Operand size Override	Operand size Override	66H	用该前缀来区分：访问 32 位或 16 位操作数；也用来区分 128 位和 64 位操作数
Address Override	Address Override	67H	64 位下指定用 64 位还是 32 位寄存器作为索引

4.9 x86机器指令编码规则

操作码

- 指明了要进行的操作
- 指明操作数的类型（一般看最后一个二进制位）

  0：字节操作； 1：字操作（32位指令中为双字操作）；

  有指令前缀 66H时，对字操作。
- 多数双操作数指令，指明源操作数是寄存器寻址，还是目的操作数是寄存器寻址 （操作码的倒数第二个二进制位）

  1：目的操作数是寄存器寻址，0：源操作数寄存器寻址

  与[寻址方式字节]配合使用
- 在有些指令中（如立即数传送给寄存器），操作码含有寄存器的编码。
- 一般在opcode的编码中体现了源操作数是否为立即数。

4.9 x86机器指令编码规则

内存/寄存器操作数(ModR/M)

Mod （6-7 位）	Reg/Opcode（3-5 位）	R/M （0-2 位）

- Mod由2个二进制位组成，取值是00、01、10、11。
- Mod与为R/M配合使用，明确一个操作的获取方法。
- Reg/Opcode 确定另外一个寄存器寻址的寄存器编码

4.9 x86机器指令编码规则

内存/寄存器操作数(ModR/M)

Mod （6-7 位）	Reg/Opcode(3-5 位)	R/M （0-2 位）

R/M	Mod
[EAX], [ECX], [EDX], [EBX], [--][--], disp32, [ESI], [EDI]	00
[EAX]+disp8, [ECX]+disp8, [EDX]+disp8, [EBX]+disp8, [--][--]+disp8, [EBP]+disp8, [ESI]+disp8, [EDI]+disp8	01
[EAX]+disp32, [ECX]+disp32, [EDX]+disp32, [EBX]+disp32, [--][--]+disp32, [EBP]+disp32, [ESI]+disp32, [EDI]+disp32	10
EAX/AX/AL/MM0/XMM0, ECX/CX/CL/MM1/XMM1, EDX/DX/DL/MM2/XMM2, EBX/BX/BL/MM3/XMM3, ESP/SP/AH/MM4/XMM4, EBP/BP/CH/MM5/XMM5, ESI/SI/DH/MM6/XMM6, EDI/DI/BH/MM7/XMM7	11

Mod=00, R/M =000，表示 用 [EAX] 寻址

Mod=00，R/M=100, 表示用[--][--] ，无位移量的基址加变址

在 SIB 字节指明基址/变址寄存器的编码

4.9 x86机器指令编码规则

内存/寄存器操作数(ModR/M)

Mod （6-7 位）	Reg/Opcode（3-5 位）	R/M （0-2 位）

000	001	010	011	100	101	110	111
AL	CL	DL	BL	AH	CH	DH	BH
AX	CX	DX	BX	SP	BP	SI	DI
EAX	ECX	EDX	EBX	ESP	EBP	ESI	EDI
MMO	MM1	MM2	MM3	MM4	MM5	MM6	MM7
XMMO	XMM1	XMM2	XMM3	XMM4	XMM5	XMM6	XMM7

Reg/Opcode 的编码

- 同一编码有多个寄存器
- 用哪一个寄存器，取决于指令前缀和操作码中的编码

4.9 x86机器指令编码规则

索引寻址描述 SIB

Scale （6-7 位）	Index(3-5 位)	Base （0-2 位）

000	001	010	011	100	101	110	111
AL	CL	DL	BL	AH	CH	DH	BH
AX	CX	DX	BX	SP	BP	SI	DI
EAX	ECX	EDX	EBX	ESP	EBP	ESI	EDI
MMO	MM1	MM2	MM3	MM4	MM5	MM6	MM7
XMMO	XMM1	XMM2	XMM3	XMM4	XMM5	XMM6	XMM7

- 与内存/寄存器操作数(ModR/M) 配合使用
- 在32位指令系统中，基址寄存器与变址寄存器都是 32 位的
- Base =000，表示基址寄存器为 EAX

4.9 x86机器指令编码规则

地址偏移量(Displacement)

- 由1、2 或 4 个字节组成，分别对应8位、16位或32位的偏移量;
- 数据按照小端顺序存放，即数据的低位存放在小地址单元中。

立即数(Immediate)

- 对应立即寻址方式
- 占1、2 或 4个字节, 按照小端顺序存放。