

计算机组成原理

Principles of Computer Organization

第 10 讲 计算机中数的运算IV

Booth乘法、快速乘法器、定点除法

主讲教师:石 侃

shikan@ict.ac.cn

2025年3月26日

④ Booth 算法递推公式

6.3

$$\begin{aligned} [x \cdot y]_{\stackrel{?}{\uparrow}h} &= [x]_{\stackrel{?}{\uparrow}h} [(y_1 - y_0) + (y_2 - y_1) 2^{-1} + \cdots + (y_{n+1} - y_n) 2^{-n}] \\ [z_0]_{\stackrel{?}{\uparrow}h} &= 0 \\ [z_1]_{\stackrel{?}{\uparrow}h} &= 2^{-1} \{ (y_{n+1} - y_n) [x]_{\stackrel{?}{\uparrow}h} + [z_0]_{\stackrel{?}{\uparrow}h} \} \qquad y_{n+1} = 0 \end{aligned}$$

:
$$[z_n]_{\not \uparrow \downarrow} = 2^{-1} \{ (y_2 - y_1)[x]_{\not \uparrow \downarrow} + [z_{n-1}]_{\not \uparrow \downarrow} \}$$

 $[x \cdot y]_{\stackrel{\text{def}}{=}} = [z_n]_{\stackrel{\text{def}}{=}} + (y_1 - y_0)[x]_{\stackrel{\text{def}}{=}}$

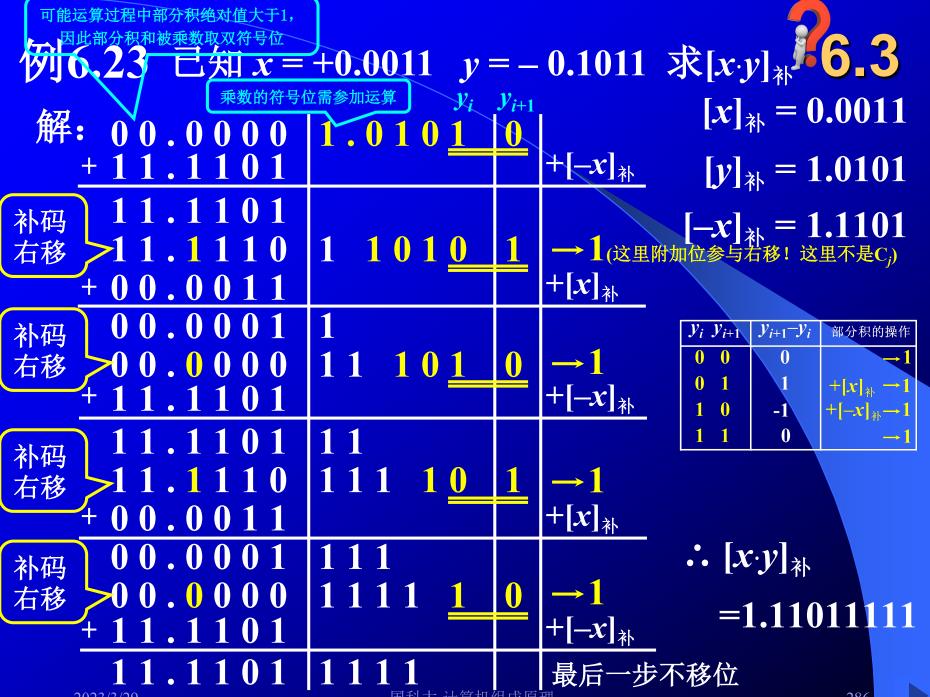
 $-[x]_{\stackrel{\wedge}{\uparrow}} = +[-x]_{\stackrel{\wedge}{\uparrow}}$

最后一步不移位

如何实现 $y_{i+1}-y_i$?

Invented by Andrew Donald Booth in 1950 while doing research on crystallography at Birkbeck College in Bloomsbury,

$y_i y_{i+1}$	y_{i+1} - y_i		部分积的操作
0 0	0		→1
0 1	1	V	$+[x]_{\lambda} \rightarrow 1$
1 0	-1	_	$+[x]_{\not \uparrow \downarrow} \rightarrow 1$ $+[-x]_{\not \uparrow \downarrow} \rightarrow 1$
1 1	0		→1



2023/3/29

国科大-计算机组成原理

286

		$y_i y_{i+1}$	$y_{i+1}-y_i$	部分积的擦
例 6.21 已知[x]**=0.1101,[y]**=0.101	l ,求[x·y]**。	0 0	0	_
可能运解:表 6.16 列出了例 6.21 的求解过程。		0 1	1	$+[x]_{i}$
算过程 中部分 乘数的符号位需参加运算 表 6.16 例 6.21 对	t[x·y]*的过程	1 0	-1	+[-x] _* +
积绝对 部分和 乘数 7 附加位		1 1	0	_
<u>値大于</u> 1, 00.0000 0101 <u>1</u> 0 分积和 + 11.0011 被乗数	初值 $[z_0]_{\#}=0$ $y_ny_{n+1}=10$,部分积	lbu[- x]	*	
取双符 号位 11.0011 11.1001 1010 <u>1</u> 11.1100 1101 <u>0</u> +00.1101	→1 位,得 $[z_i]_{*}$ $y_n y_{n+1} = 11$,部分形 $y_n y_{n+1} = 01$,部分形		导[z ₂] _料	
00.1001 11 00.0100 1110 <u>1</u> +11.0011	→I 位,得[z ₃] _計 y _n y _{n+1} = 10,部分利	₹加[- *]	3 1	
11.0111 111 11.1011 1111 <u>0</u> +00.1101	y,y,+,=01,部分和			
00.1000 1111	最后一步不移位,	得[x·y]	#	

故 $[x \cdot y]_{\#} = 0.10001111$ 2023/3/29

例 6.22	已知[x]*	= 1.0101	$[\gamma]_{*}$	= 1.0011	,求[x
--------	--------	----------	----------------	----------	------

$y_i y_{i+1}$	$y_{i+1}-y_i$	部分积的操作
0 0	0	→1
0 1	1	$+[x]_{i}\rightarrow 1$
1 0	-1	$+[-x]_{\uparrow\downarrow} \rightarrow 1$
1 1	0	→1

部分积	乘数 %。	附加位 y _{n+1}	说明
00.0000	10011	<u>ō</u>	
+00.1011			y,y,,, = 10,部分积加[-x],
0.0.1011			
00.0101	11001	1	→1 位,得[z ₁]**
00.0010	11100	1	$y_n y_{n+1} = 11$, 部分积→1 位, 得 $[z_2]_*$
+ 11.0101			y, y, y; = 01, 部分积加[x];h
11.0111	1 1		
11.1011	11110	ō	→1 位,得[z ₃] ₊
11.1101	11111	ō	$y_n y_{n+1} = 00$,部分 $\rightarrow 1$ 位,得 $\{z_4\}_{*}$
+00.1011			y,y,+,=10,部分积加[-x]**
00.1000	1111		最后一步不移位,得[x·y]*

故 [x·y]** = 0.10001111

Long Table → Short Table

算的省心(符号任意)、算的快(某些情况仅做右移操作)

【这里的P即前述的部分积z,另外这里yBooth布斯算法举例的脚标序号是递减的,前述是递增的】

已知[X]_补 = 1_0110_1101,[Y]_补 = 0_1111_1110,计算[X×Y]_补 $[-X]_{3/2} = 0_1001_0011 X=-147, Y=254, X\times Y=-37338,$



[X×Y]_补应等于1_0110_1110_0010_0110 验证: [Y]_补中有连续1时,加速明显

''			••
Р	Υ	y_{-1}	说明
00_0000_0000	0_1111_111 <u>0</u>	<u>0</u>	y_1为0,[P ₀] _补 为0
00_0000_0000	0 0111_111 <u>1</u>	<u>0</u>	$y_0y_{-1} = 00$, PY右移一位
+ 00 1001 0011 00 1001 0011 00 0100 1001	10 011_111 <u>1</u>	<u>1</u>	$y_1y_0 = 10$, $+[-x]_{\dot{\gamma} }$, PY右移一位
00_0010_0100	110 01_111 <u>1</u>	<u>1</u>	y ₂ y ₁ = 11, P Y直接右移一位
00_0001_0010	0110 0_111 <u>1</u>	<u>1</u>	$y_3y_2=11$, PY直接右移一位
00_0000_1001	0_0110 011 <u>1</u>	<u>1</u>	$y_4y_3=11$, PY直接右移一位
00_0000_0100	10_0110 01 <u>1</u>	<u>1</u>	$y_5y_4=11$, PY直接右移一位
00_0000_0010	010_0110 0 <u>1</u>	<u>1</u>	$y_6y_5=11$, PY直接右移一位
00_0000_0001	0010_0110 <u>0</u>	<u>1</u>	$y_7y_6=11$, PY直接右移一位
+ 11_0110_1101 11_0110_1110	0010_0110		$y_8y_7 = 01, +[x]_{\stackrel{?}{\uparrow}},$ 最后一步不移位

Booth算法的实质和优点

 y_i y_{i+1} y_{i+1} y_i 部分积的操作

 0
 0
 \rightarrow 1

 0
 1
 1
 $+[x]_{\uparrow}$ \rightarrow 1

 1
 0
 -1 $+[-x]_{\uparrow}$ \rightarrow 1

 1
 1
 0
 \rightarrow 1

乘数的一轮移位操作:

middle of run

0 1 1 1 0

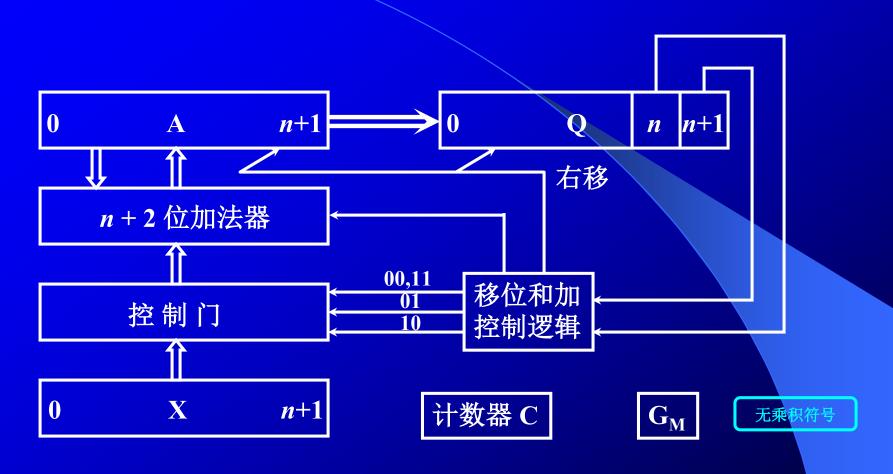
d of run beginning of run

◆ 当前位 <i>y_i</i>	右边位 y_{i+1}	除移位外的操作	Example
1	0	减被乘数	000111 <u>10</u> 00
1	1	加0 (不操作)	00011 <u>11</u> 000
0	1	加被乘数	00 <u>01</u> 111000
0	0	加0 (不操作)	0 <u>00</u> 1111000

- ◆在"1串"中,第一个1时做减法,最后一个1做加法,其余情况只要移位
- ◆同前面算法一样,将乘积寄存器右移一位。(这里是算术右移)
- ◆假设一个8位乘数: 0111_1110, 它将产生6行非零的部分积。如果把该数字记成另一种形式,(1)000_00(-1)0(-1是负1),则只需相加两个部分积,可以大大减少非零行的数目,意味着相加次数的减少,从而加快了运算速度
- ◆ Booth算法可以减少部分积的数目,用来计算有符号乘法,提高乘法速度

(2) Booth 算法的硬件配置

6.3



A、X、Q均为n+2位寄存器(A与X为双符号位,Q为单符号位+附加位)移位和加操作受乘数的末两位控制

补充知识: 补码两位乘法

◆ **补码两位乘可用布斯算法推导如下**:【这里的P即前述的部分积z,另外这里y的脚标序号是递减的,前述是递增的】

•
$$[P_{i+1}]_{?|} = 2^{-1} ([P_i]_{?|} + (y_{i-1} - y_i) [X]_{?|})$$

• $[P_{i+2}]_{?|} = 2^{-1} ([P_{i+1}]_{?|} + (y_i - y_{i+1}) [X]_{?|})$
= $2^{-1} (2^{-1} ([P_i]_{?|} + (y_{i-1} - y_i) [X]_{?|}) + (y_i - y_{i+1}) [X]_{?|})$
= $2^{-2} ([P_i]_{?|} + (y_{i-1} + y_i - 2y_{i+1}) [X]_{?|})$

- ◆ 开始置附加位y₋₁为0,乘积寄存器最高位前面添加一位附加符号位0。
- ◆ 最终的乘积高位部分在乘积寄存器P中,低位部分在乘数寄存器Y中。
- ◆ 因为字长总是8的倍数,所以 补码的位数n应该是偶数,因 此,总循环次数为n/2。

y _{i+1}	y _i	y _{i-1}	操作	迭代公式
0	0	0	0	2-2[P _i] _*
0	0	1	+[X] _{ネト}	2 ⁻² {[P _i] _补 +[X] _补 }
0	1	0	+[X] _{ネト}	$2^{-2}\{[P_i]_{i}+[X]_{i}\}$
0	1	1	+2[X] _{ネト}	$2^{-2}\{[P_i]_{i}+2[X]_{i}\}$
1	0	0	+2[-X] _补	2 ⁻² {[P _i] _补 +2[-X] _补 }
1	0	1	+[-X] _{ネト}	2 ⁻² {[P _i] _补 +[-X] _补 }
1	1	0	+[-X] _{ネト}	2 ⁻² {[P _i] _补 +[-X] _补 }
1	1	1	0	2 ⁻² [P _i] _*

- 1. 原码与补码的乘法,根本区别在于对符号位的处理
- 2. 补码(一位/两位)乘法:符号和数值一起运算,运算结果的符号是在数值部分的运算过程中自然形成
 - ① 补码一位乘、乘数为负数时,校正法(加[-x]*)
 - ②补码一位乘、乘数正负任意,比较法(Booth算法)
- 3. 注意区别 $[-x^*]_{*}$ 和 $[-x]_{*}$
 - ① 原码两位乘法中使用[-x*]*
 - ② 补码乘法中使用[-x]补

y_i .	y_{i+1}	$y_{i+1} - y_i$	部分积的操作
0	0	0	$\rightarrow 1$
0	1	1	$+[x]_{\uparrow \downarrow} \rightarrow 1$
1	0	-1	$+[-x]_{\nmid h} \rightarrow 1$
1	1	0	→1

4. 由于不同的机器数运算规则不同,运算器硬件组成各 不相同(包括寄存器位数、全加器输入端控制电路等)

快速乘法器

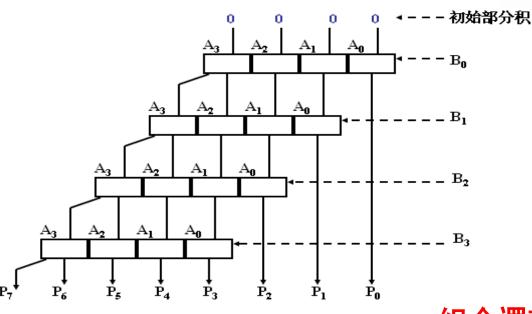
- ◆前面介绍的乘法部件的特点
 - 通过一个ALU多次做"加/减+右移"来实现
 - 一位乘法:约n次"加+右移"
 - 两位乘法:约n/2次"加+右移"

所需时间随位数增多而加长,由时钟和控制电路控制

- ◆设计快速乘法部件的必要性
 - 大约1/3是乘法运算
- ◆快速乘法器的实现(由特定功能的组合逻辑单元构成)
 - 流水线方式/硬件叠加方式(如: 阵列乘法器, 附录6B)
 - 所有部分积并行相加,组织为二叉树结构,例如: $16\rightarrow 8\rightarrow 4\rightarrow 2\rightarrow$ 结果
 - Booth算法(Andrew Donald Booth, 1950) + Wallace Tree (Chris Wallace, 1964)
 - 华莱士树(Wallace Tree):硬件快速把n个数相加归约为2个数的相加,从而加速多个部分积相加速度【体系结构课程会继续介绍】

用"空间"换"时间"

流水线方式的快速乘法器



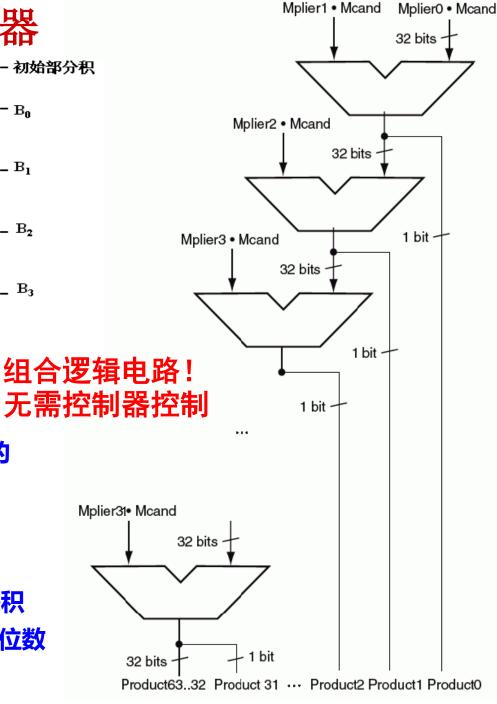
◆ 为乘数的每位提供一个n位加法器

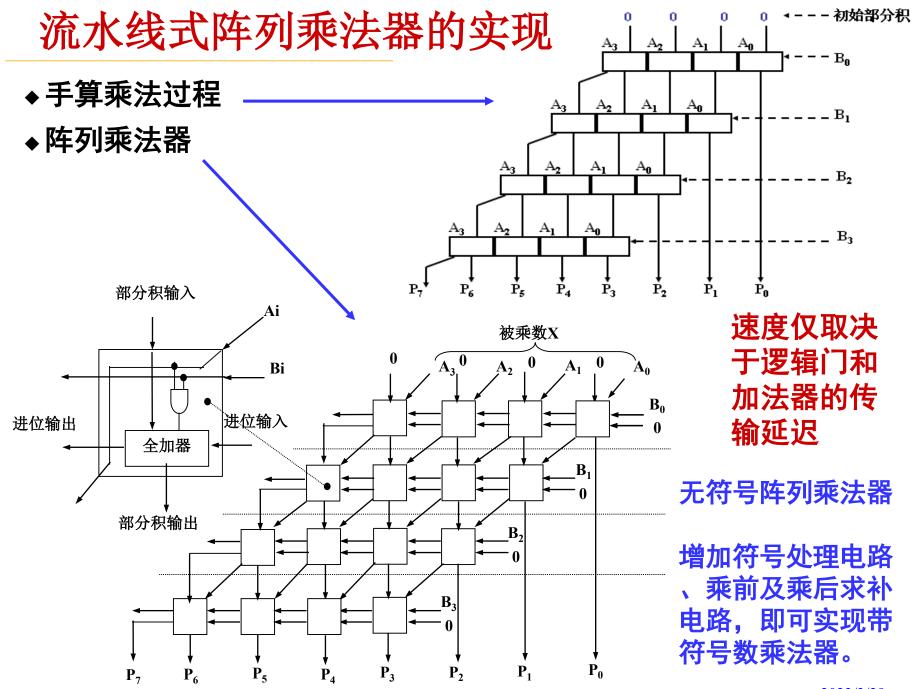
◆每个加法器的两个输入端分别是:

・本次乘数对应的位与被乘数相与的

结果 (即: 0或被乘数)

- 上次部分积
- ◆每个加法器的输出分为两部分:
 - ·和的最低有效位(LSB)作为本位乘积
 - · 进位和高31位的和数组成一个32位数 作为本次部分积





Google Tensor Processing Unit





(paper released on Apr 5, 2017)

In-Datacenter Performance Analysis of a Tensor Processing Unit

Norman P. Jouppi, Cliff Young, Nishant Patil, David Patterson, Gaurav Agrawal, Raminder Bajwa, Sarah Bates, Suresh Bhatia, Nan Boden, Al Borchers, Rick Boyle, Pierre-luc Cantin, Clifford Chao, Chris Clark, Jeremy Coriell, Mike Daley, Matt Dau, Jeffrey Dean, Ben Gelb, Tara Vazir Ghaemmaghami, Rajendra Gottipati, William Gulland, Robert Hagmann, C. Richard Ho, Doug Hogberg, John Hu, Robert Hundt, Dan Hurt, Julian Ibarz, Aaron Jaffey, Alek Jaworski, Alexander Kaplan, Harshit Khaitan, Daniel Killebrew, Andy Koch, Naveen Kumar, Steve Lacy, James Laudon, James Law, Diemthu Le, Chris Leary, Zhuyuan Liu, Kyle Lucke, Alan Lundin, Gordon MacKean, Adriana Maggiore, Maire Mahony, Kieran Miller, Rahul Nagarajan, Ravi Narayanaswami, Ray Ni, Kathy Nix, Thomas Norrie, Mark Omernick, Narayana Penukonda, Andy Phelps, Jonathan Ross, Matt Ross, Amir Salek, Emad Samadiani, Chris Severn, Gregory Sizikov, Matthew Snelham, Jed Souter, Dan Steinberg, Andy Swing, Mercedes Tan, Gregory Thorson, Bo Tian, Horia Toma, Erick Tuttle, Vijay Vasudevan, Richard Walter, Walter Wang, Eric Wilcox, and Doe Hyun Yoon Google, Inc., Mountain View, CA USA

jouppi@google.com



Figure 3. TPU Printed Circuit Board. It can be inserted in the slot for an SATA disk in a server, but the card uses PCIe Gen3 x16.

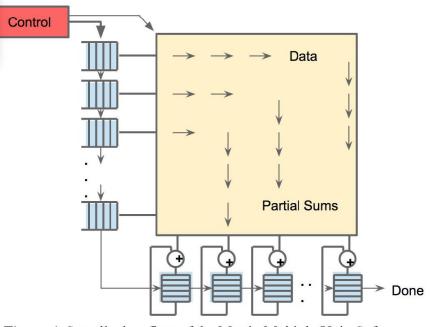


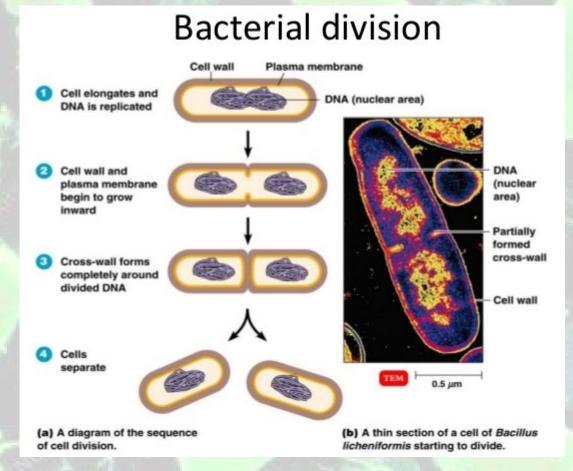
Figure 4. Systolic data flow of the Matrix Multiply Unit. Software has the illusion that each 256B input is read at once, and they instantly update one location of each of 256 accumulator RAMs.

Norman P. Jouppi et al., In-Datacenter Performance Analysis of a Tensor Processing Unit, ISCA 2017

乘法小结

- 整数乘法与小数乘法完全相同可用 逗号 代替小数点
- ▶ 原码乘 符号位 单独处理 补码乘 符号位 自然形成
- > 原码乘去掉符号位运算 即为无符号数乘法
- > 不同的乘法运算需有不同的硬件支持

Why are bacteria **bad** at math? Because they **multiply** by dividing.



四、除法运算

6.3

1. 分析笔算除法

$$x = -0.1011$$
 $y = 0.1101$ $\Re x \div y$

0.1101 (商)

0.1101) 0.10110

(除数)

这里都是余数的绝对

值。最后一步得到的 是真正余数的绝对值

R*,而R的实际符号位 应与被除数一致,即

从被除数直接获得,

所以真正余数应为

-0.00000111

0.10110

(被除数)

0.01101(右移的除数)

0.010010(中间余数)

0.001101

0.00010100

0.00001101

0.00000111(余数)

✓商符单独处理

- ? 心算上商(是否够减,被除 数或中间余数和除数比大小)
- ?余数不动低位补"0"减右移一位的除数
- ? 上商位置不固定(商是 从高位向低位逐位求得)

 $x \div y = -0.1101$

商符心算求得

余数 0.00001111

这是余数的绝对值R*

真正余数R = - 0.00000111

2. 笔算除法和机器除法的比较

6.3

笔算除法

商符单独处理 心算上商

余数 不动 低位补 "0" 减右移一位 的除数

2 倍字长加法器

上商位置 不固定

机器除法

被除数与除数符号位异或

|x|-|y|>0(够减)上商1

| x | − | y | < 0 (不够减) 上商 0

余数左移一位低位补"0"减除数

1 倍字长加法器

在寄存器 最末位上商,并且部分商左移一位

3. 原码除法

6.3

以小数为例

$$[x]_{\mathbb{R}} = x_0. x_1 x_2 \dots x_n$$

$$[y]_{\mathbb{R}} = y_0. y_1 y_2 \dots y_n$$

$$[\frac{x}{y}]_{\mathbb{R}} = (x_0 \oplus y_0). \frac{x^*}{y^*}$$

式中
$$x^* = 0. x_1 x_2 \cdots x_n$$
 为 x 的绝对值 $y^* = 0. y_1 y_2 \cdots y_n$ 为 y 的绝对值

商的符号位单独处理 $x_0 \oplus y_0$

数值部分为绝对值相除



约定

小数定点除法 x* < y*
被除数不等于 0
除数不能为 0
(在除法运算之前先做检查)

整数定点除法 x* > y* 商的位数与其他操作数位数相同(位数达到一致则停止运算)

(1) 恢复余数法

6.3

例6.24 x = -0.1011 y = -0.1101 求 $\left[\frac{x}{y}\right]_{\mathbb{R}}$

解: $[x]_{\mathbb{R}} = 1.1011$ $[y]_{\mathbb{R}} = 1.1101$ $[y^*]_{\mathbb{A}} = 0.1101$ $[-y^*]_{\mathbb{A}} = 1.0011$

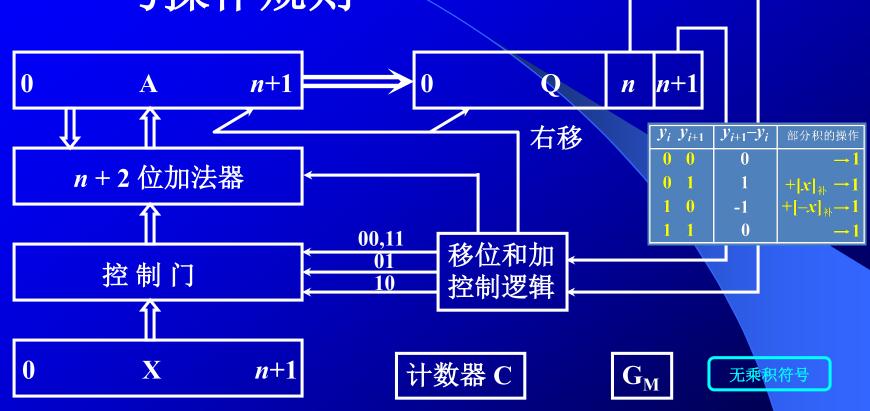
② 被	除数(余数)	商	说明	
0	.1011	0.0000		
+ 1	.0011		+[- y*]* (进行x*-y*作	[比较]
1	.1110	0	余数为负,上商 0	
+ 0	.1101		恢复余数 +[y*] _补	
() WHEH - 150	.1011	0	恢复后的余数	
逻辑左移 1	.0110	0	←1(余数和商同时左	移)
为什么是 + 1 逻辑左移?	.0011		+[-y*] _补	
	.1001	0 1	余数为正,上商1	
逻辑左移 1	.0010	0 1	<u>←</u> 1	
+ 1	.0011		+[- <i>y*</i>] _补	

2023/3/29

国科人-订昇机组成原理





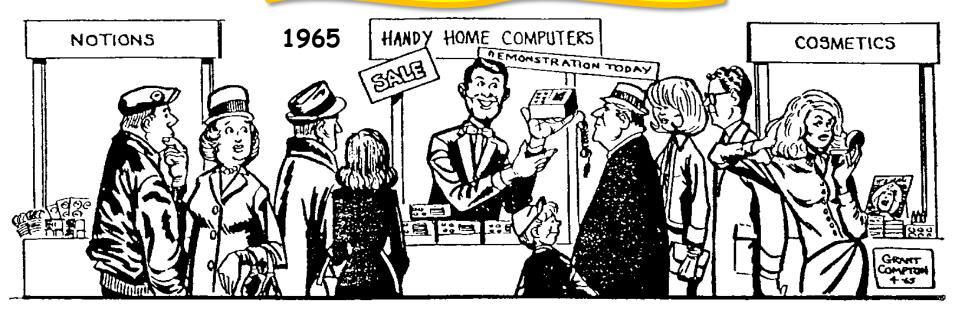


A、X、Q均为n+2位寄存器_(A与X为双符号位,Q为单符号位+附加位)移位和加操作受乘数的末两位控制

Best sellers in the **BOOTH**

2023

ChatGPT、GPT-4、Bard等大模型支撑下的 AIGC服务及其底层的软硬件系统与AI芯片





"Generative AI is a new kind of computer, one we program in human language.

Now, everyone is a programmer"

---Jensen Huang, NVIDIA's Spring GTC 2023 Keynote

AIGC: Artificial Intelligence Generated Content
https://openai.com/research/gpt-4
https://bard.google.com
https://www.anandtech.com/show/18782/the-nvidia-gtc-spring-2023-keynote-live-blog-800am-pt1500-utc

Moore's Law Predicts: 2X Transistors / chip every 2 years

The experts look ahead

Cramming more components onto integrated circuits

With unit cost falling as the number of components per circuit rises, by 1975 economics may dictate squeezing as many as 65,000 components on a single silicon chip

By Gordon E. Moore

Director, Research and Development Laboratories, Fairchild Semiconductor division of Fairchild Camera and Instrument Corp.

The future of integrated electronics is the future of electronics itself. The advantages of integration will bring about a proliferation of electronics, pushing this science into many

Integrated circuits will lead to such wonders as home computers or at least terminals connected to a central computer-automatic controls for automobiles, and personal portable communications equipment. The electronic wristwatch needs only a display to be feasible today.

But the biggest potential lies in the production of large systems. In telephone communications, integrated circuits in digital filters will separate channels on multiplex equipment. Integrated circuits will also switch telephone circuits and perform data processing.

Committees will be more powerful, and will be organized in completely different ways. For example, memories built of integrated electronics may be distributed throughout the

Th	Ca
The author	be
Dr. Gordon E. Moore is one of	01
the new breed of electronic	
engineers, schooled in the	
physical sciences rather than in	ad
electronics. He earned a B.S.	to
degree in chemistry from the	to
University of California and a	GI
Ph.D. degree in physical	
chemistry from the California	131
Institute of Technology. He was	ra
one of the founders of Fairchild	
Semiconductor and has been	in
director of the research and	"
development laboratories since	
1959.	

Electronics, Volume 38, Num

Integrated electronics is established today. Its techniques are almost mandatory for new military systems, since the reliability, size and weight required by some of them is achievable only with integration. Such programs as Apollo, for manned moon flight, have demonstrated the reliability of integrated electronics by showing that complete circuit functions are as free from failure as the best individual transis-

Most companies in the commercial computer field have machines in design or in early production employing integrated electronics. These machines cost less and perform better than those which use "conventional" electronics.

Instruments of various sorts, especially the rapidly increasing numbers employing digital techniques, are starting to use integration because it cuts costs of both manufacture

The use of linear integrated circuitry is still restricted primarily to the military. Such integrated functions are expensive and not available in the variety required to satisfy a major fraction of linear electronics. But the first applications are beginning to appear in commercial electronics, particularly in equipment which needs low-frequency amplifiers of small size.

Reliability counts

In almost every case, integrated electronics has demonstrated high reliability. Even at the present level of production-low compared to that of discrete components-it offers reduced systems cost, and in many systems improved performance has been realized.

Integrated electronics will make electronic techniques more generally available throughout all of society, performing many functions that presently are done inadequately by other techniques or not done at all. The principal advantages will be lower costs and greatly simplified design-payoffs from a ready supply of low-cost functional packages.

For most applications, semiconductor integrated circuits will predominate. Semiconductor devices are the only reasonable candidates presently in existence for the active elements of integrated circuits. Passive semiconductor elements look attractive too, because of their potential for low cost and high reliability, but they can be used only if precision is

Silicon is likely to remain the basic material, although others will be of use in specific applications. For example, gallium arsenide will be important in integrated microwave functions. But silicon will predominate at lower frequencies because of the technology which has already evolved around it and its oxide, and because it is an abundant and relatively inexpensive starting material

Costs and curves

Reduced cost is one of the big attractions of integrated electronics, and the cost advantage continues to increase as the technology evolves toward the production of larger and larger circuit functions on a single semiconductor substrate For simple circuits, the cost per component is nearly inversely oportional to the number components, the result of the

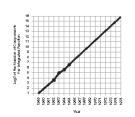


a two-mil square can also contain several kilohms of resistance or a few diodes. This allows at least 500 components per linear inch or a quarter million per square inch. Thus, 65,000 components need occupy only about one-fourth a

On the silicon wafer currently used, usually an inch or more in diameter, there is ample room for such a structure if the components can be closely packed with no space wasted for interconnection patterns. This is realistic, since efforts to achieve a level of complexity above the presently available integrated circuits are already underway using multilayer metalization patterns separated by dielectric films. Such a density of components can be achieved by present optical techniques and does not require the more exotic techniques such as electron beam operations, which are being studied to make even smaller structures

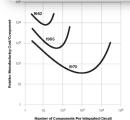
Increasing the yield

There is no fundamental obstacle to achieving device yields of 100%. At present, packaging costs so far exceed the cost of the semiconductor structure itself that there is no incentive to improve yields, but they can be raised as high as



Page 3

the e



is economically justified. No barrier exists comparable to the thermodynamic equilibrium considerations that often limit yields in chemical reactions; it is not even necessary to do any fundamental research or to replace present processes. Only the engineering effort is needed

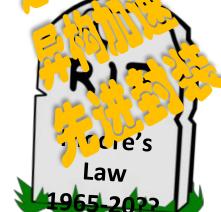


Gordon Moore Intel Co-founder 1929/1/3-2023/3/24

> rated stages of gain, giving high performance at ost, interspersed with relatively large tuning ele-

> near functions will be changed considerably. The nd tracking of similar components in integrated ill allow the design of differential amplifiers of oved performance. The use of thermal feedback abilize integrated structures to a small fraction of Il allow the construction of oscillators with crys-

the microwave area, structures included in the f integrated electronics will become increasingly The ability to make and assemble components ared with the wave engths involved will allow rameter design, at least at the lower freo predict at the present time just how of the m crowave area by integrated essful realization of such items example, usine a multiplicity



Electronics, Volume 38, Number 8, April 19, 1965

2023/3/29

Electronics, Volume 38, Number 8, April 19, 1965