

2008年3月15日  
星期六 (1)

# CEG5010: The Symmetric Table Addition Method

2008年3月15日  
星期六 (2)

## Introduction

- Lookup tables can be used to implement any function
- The Symmetric Table Addition Method (STAM) can be used to approximate any differentiable function
- In STAM, symmetry and leading zeros are introduced to reduce amount of memory needed (Stine and Schulte 1998)

2008年3月15日  
星期六 (3)

## Taylor Series

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)(x-a)^2}{2!} + \dots$$

$$+ \frac{f^{(n)}(a)(x-a)^n}{n!} + R_n$$

$$= \sum_{i=0}^n \frac{f^{(i)}(a)(x-a)^i}{i!} + R_n$$

where

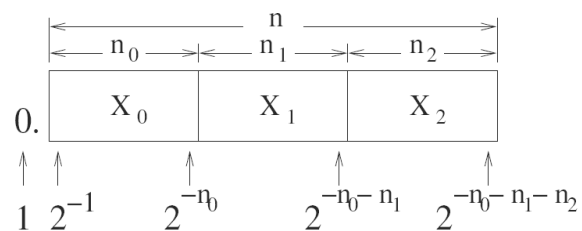
$$R_n = \int_a^x f^{(n+1)}(u) \frac{(x-u)^n}{n!} du$$

$$= \frac{f^{(n+1)}(\xi)(x-a)^{n+1}}{(n+1)!} \quad \text{for } a < \xi < x$$

2008年3月15日  
星期六 (4)

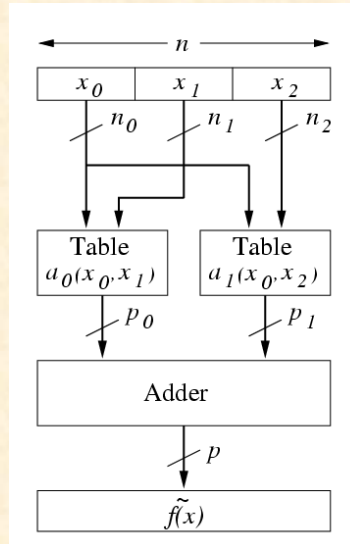
## Bipartite Table Method

- Operand  $x$  divided into 3 parts  $x = x_0 + x_1 + x_2$



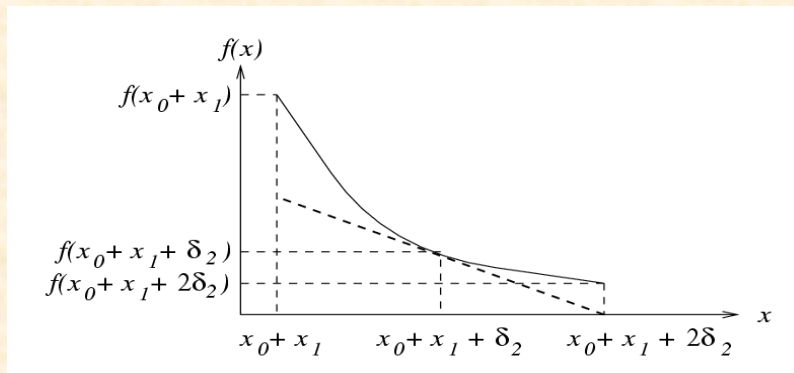
2008年3月15日  
星期六 (5)

## Bipartite Method (BM)

2008年3月15日  
星期六 (6)

## BM coefficients

- Computed by taking first 2 terms of Taylor series around  $x_0+x_1+\delta_2$  ( $\delta_2$  is half the range of  $x_2$ )



2008年3月15日  
星期六 (7)

## BM coefficients

$$\begin{aligned}
 f(x) &= f(x_0 + x_1 + x_2) \\
 &\approx f(x_0 + x_1 + \delta_2) + f'(x_0 + x_1 + \delta_2)(x_2 - \delta_2) \\
 &\approx f(x_0 + x_1 + \delta_2) + f'(x_0 + \delta_1 + \delta_2)(x_2 - \delta_2) \\
 &= a_0(x_0, x_1) + a_1(x_0, x_2) \\
 \delta_1 &= (2^{-n_0} - 2^{-n_0-n_1})/2 \\
 &= 2^{-n_0-1} - 2^{-n_0-n_1-1} \\
 \delta_2 &= (2^{-n_0-n_1} - 2^{-n_0-n_1-n_2})/2 \\
 &= 2^{-n_0-n_1-1} - 2^{-n_0-n_1-n_2-1}
 \end{aligned}$$

2008年3月15日  
星期六 (8)

## BM coefficients

- Note that  $a_1(x_0, x_2)$  cannot depend on  $x_1$  and is replaced by  $\delta_1$ , a value which is half the range of  $x_1$

$$\begin{aligned}
 f(x) &= f(x_0 + x_1 + x_2) \\
 &\approx f(x_0 + x_1 + \delta_2) + f'(x_0 + x_1 + \delta_2)(x_2 - \delta_2) \\
 &\approx f(x_0 + x_1 + \delta_2) + f'(x_0 + \delta_1 + \delta_2)(x_2 - \delta_2) \\
 &= a_0(x_0, x_1) + a_1(x_0, x_2)
 \end{aligned}$$

2008年3月15日  
星期六 (9)

## Symmetry in $a_1(x_0, x_2)$

- There is a bound on  $a_1$

$$|a_1(x_0, x_2)| = |f'(x_0 + \delta_1 + \delta_2)(x_2 - \delta_2)|$$

$$< |f'(\xi_1)| 2^{-n_0 - n_1 - 1}$$

- Leading to the number of leading zeros to be approximately

$$n_0 + n_1 + 1 + \log_2(|f(\xi_0)/f'(\xi_1)|)$$

(the leading zeros do not need to be stored)

2008年3月15日  
星期六 (10)

## Symmetry in $a_1(x_0, x_2)$

- By symmetry,  $a_1(x_1, x_2)$  can be reduced in size from  $2^{n_0+n_2} \times p_1$  bits to  $2^{n_0+n_2-1} \times (p_1 - 1)$  bits

$x$	$x_2 2^7$	$a_1(x_0, x_2)$	
		decimal	binary
0.500000	000	+0.0166016	0.0000010001
0.507812	001	+0.0107422	0.0000001011
0.515625	010	+0.0068359	0.0000000111
0.523438	011	+0.0029297	0.0000000011
0.531250	100	-0.0029297	1.1111111101
0.539062	101	-0.0068359	1.1111111001
0.546875	110	-0.0107422	1.1111110101
0.554688	111	-0.0166016	1.1111101111

$$f(x) = \cos(x)$$

$$x_0 = 0.5$$

$$x_1 = 0$$

$$n_0 = 2$$

$$n_1 = 2$$

$$n_2 = 3$$

$$p = 7$$

$$g = 2$$

**Table 1. Table Entries of  $a_1(x_0, x_2)$ .**

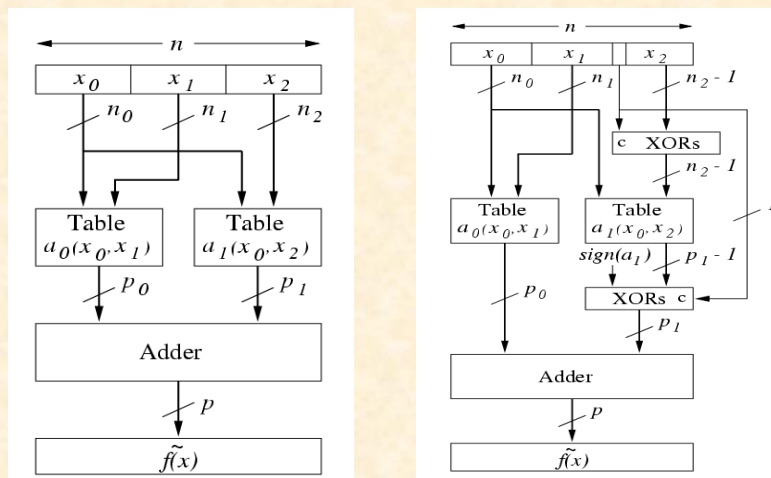
2008年3月15日  
星期六 (11)

# Symmetry

- Table is reduced by using
  - $2\delta_2 - x_2$  is the one's complement of  $x_2$
  - $a_1(x_0, 2\delta_2 - x_2)$  is the one's complement of  $a_1(x_0, x_2)$

2008年3月15日  
星期六 (12)

# Symmetric Bipartite Table Method (SBTM)



2008年3月15日  
星期六 (13)

## 4 Error sources

- $\epsilon_0$ : approximating  $f(x)$  with only the first 2 terms of the Taylor series

$$\begin{aligned}\epsilon_0 &= \left| \sum_{i=2}^{\infty} \frac{f^{(i)}(x_0 + x_1 + \delta_2)(x_2 - \delta_2)^i}{i!} \right| \\ &\approx \frac{|f''(x_0 + x_1 + \delta_2)|(x_2 - \delta_2)^2}{2!} \\ &< |f''(\xi_2)|2^{-2n_0-2n_1-2}\end{aligned}$$

$\xi_i$  is the point on  $[0,1)$  at which  $|f^{(i)}(x)|$  is max

2008年3月15日  
星期六 (14)

## 4 Error sources

- $\epsilon_1$ : replacing the 2<sup>nd</sup> term of the Taylor series with  $a_1(x_0, x_2)$

$$\begin{aligned}\epsilon_1 &\approx |f''(x_0 + \delta_1 + \delta_2)(x_1 - \delta_1)(x_2 - \delta_2)| \\ &< |f''(\xi_2)|2^{-2n_0-n_1-2}\end{aligned}$$

2008年3月15日  
星期六 (15)

## 4 Error sources

- $\epsilon_2$ : rounding  $a_0(x_0, x_1)$  and  $a_1(x_0, x_2)$  to  $p_f + g$  fractional bits

$$\begin{aligned}\epsilon_2 &\leq 2 \times 2^{-p_f - g - 1} \\ &= 2^{-p_f - g}\end{aligned}$$

2008年3月15日  
星期六 (16)

## 4 Error sources

- $\epsilon_3$ : rounding to the final result

$$\epsilon_3 \leq 2^{-p_f - 1} - 2^{-p_f - g - 1}$$

2008年3月15日  
星期六 (17)

## Error analysis

- Assume  $p_f$  fraction bits and  $p_f+g$  coefficient bits, need to ensure that the sum of these errors is smaller than 1 unit in the last place (ulp)

$$\epsilon_0 + \epsilon_1 + \epsilon_2 + \epsilon_3 \leq 2^{-p_f}$$

- This leads to the constraint

$$|f''(\xi_2)|2^{-2n_0-n_1-2}(1+2^{-n_1})+2^{-p_f-g-1} \leq 2^{-p_f-1}$$

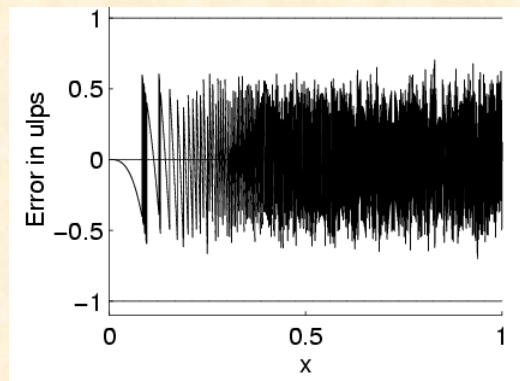
- Satisfied by

$$\begin{aligned} g &= 2 \\ 2n_0 + n_1 &\leq p_f + \log_2(f''(\xi_2)) \end{aligned}$$

2008年3月15日  
星期六 (18)

## Example

- Approximate  $\sin(x)$  to  $p_f=12$  bits on  $[0,1)$
- $n_0=4, n_1=4, n_2=4$
- $2 \times 4 + 4 \geq 12 + \lceil \log_2(1/\sqrt{2}) \rceil$



2008年3月15日  
星期六 (19)

# Examples

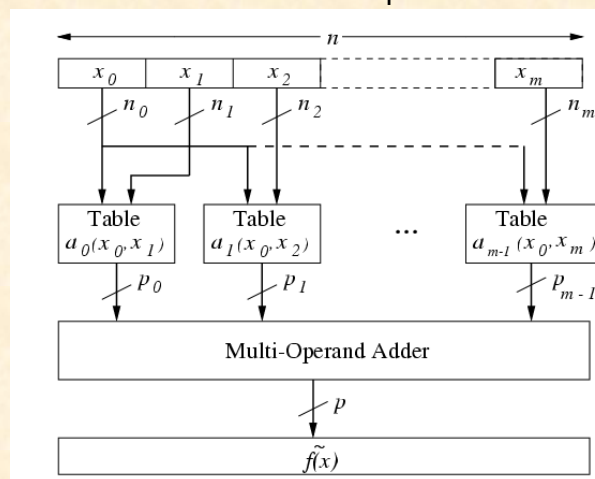
Table 2. Ranges and Derivative Values for Selected Functions.

$f(x)$	Input	Output	$ f'(\xi_1) $	$ f''(\xi_2) $
$1/x$	[1, 2)	(0.5, 1]	1	2
$\sqrt{x}$	[1, 2)	$[1, \sqrt{2})$	1/2	1/4
$\sin(x)$	[0, 1)	$[0, 1/\sqrt{2})$	1	$1/\sqrt{2}$
$2^x$	[0, 1)	[1, 2)	$2 \ln 2$	$2(\ln(2))^2$
$\log_2(x)$	[1, 2)	[0, 1)	$\log_2(e)$	$\log_2(e)$
$\ln(x)$	[1, 2)	$[0, \ln(2))$	1	1
$1/\sqrt{x}$	[1, 2)	$(1/\sqrt{2}, 1]$	1/2	3/4

2008年3月15日  
星期六 (20)

# Symmetric Table Addition Method

- Generalization of STAM with multiple tables



2008年3月15日  
星期六 (21)

## STAM approximation

$$\begin{aligned}\widetilde{f(x)} &= f(x_0 + x_1 + \sum_2^m \delta_i) + \\ &\quad f'(x_0 + \delta_1 + \sum_2^m \delta_i) \left( \sum_2^m x_i - \sum_2^m \delta_i \right) \\ &= a_0(x_0, x_1) + \sum_2^m a_{i-1}(x_0, x_i)\end{aligned}$$

$$a_{i-1}(x_0, x_i) = f'(x_0 + \delta_1 + \sum_2^m \delta_k)(x_i - \delta_i) \quad 2 \leq i \leq m$$

2008年3月15日  
星期六 (22)

## STAM error constraints

$$\begin{aligned}2n_0 + n_1 &\leq p_f + \log_2(|f''(\xi_2)|) \\ g &\leq 2 + \log_2(m - 1)\end{aligned}$$

2008年3月15日  
星期六 (23)

## Partitions and table sizes (16 bit)

Table 3. Bit Partitions and Table Sizes for 16-Bit Operands.

$f(x)$	Tables	Bit Partitions	Table Sizes	Total Memory
$1/x$	1	15	$2^{15} \cdot 15$	491,520
$1/x$	2	6, 4, 5	$2^{10}(17 + 7)$	24,576
$1/x$	3	7, 2, 3, 3	$2^9(18 + 9 + 6)$	16,896
$1/x$	4	7, 2, 2, 2, 2	$2^9 \cdot 19 + 2^8(10 + 8 + 6)$	15,872
$\sqrt{x}$	1	15	$2^{15} \cdot 15$	491,520
$\sqrt{x}$	2	4, 5, 6	$2^9(17 + 7)$	12,288
$\sqrt{x}$	3	5, 3, 3, 4	$2^8 \cdot 18 + 2^7 \cdot 9 + 2^8 \cdot 6$	7,296
$\sqrt{x}$	4	5, 3, 2, 2, 3	$2^8 \cdot 19 + 2^6(10 + 8) + 2^7 \cdot 6$	6,784
$\sin(x)$	1	16	$2^{16} \cdot 16$	1,048,576
$\sin(x)$	2	6, 4, 6	$2^{10} \cdot 18 + 2^{11} \cdot 7$	32,768
$\sin(x)$	3	7, 2, 3, 4	$2^9(19 + 9) + 2^{10} \cdot 6$	20,480
$\sin(x)$	4	7, 2, 2, 2, 3	$2^9 \cdot 20 + 2^8(10 + 8) + 2^9 \cdot 6$	17,920
$2^x$	1	16	$2^{16} \cdot 15$	983,040
$2^x$	2	5, 5, 6	$2^{10}(17 + 7)$	24,576
$2^x$	3	6, 3, 3, 4	$2^9 \cdot 18 + 2^8 \cdot 9 + 2^9 \cdot 6$	14,592
$2^x$	4	6, 3, 2, 2, 3	$2^9 \cdot 19 + 2^7(10 + 8) + 2^8 \cdot 6$	13,568

2008年3月15日  
星期六 (24)

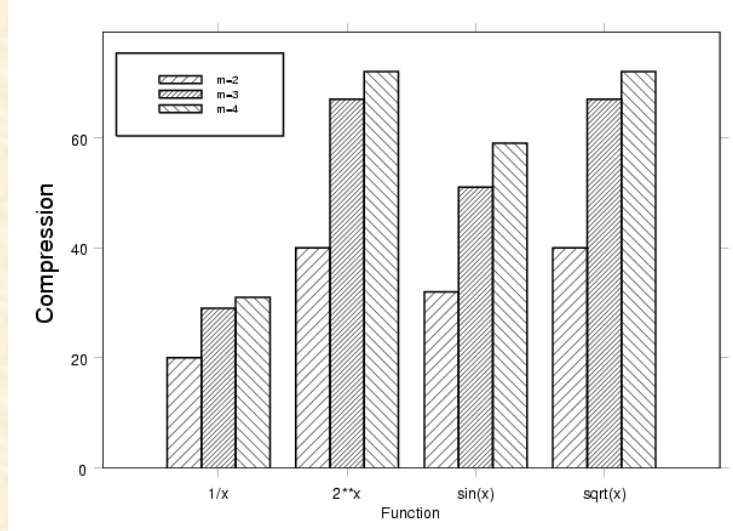
## Partitions and table sizes (24 bit)

Table 4. Bit Partitions and Table Sizes for 24-Bit Operands.

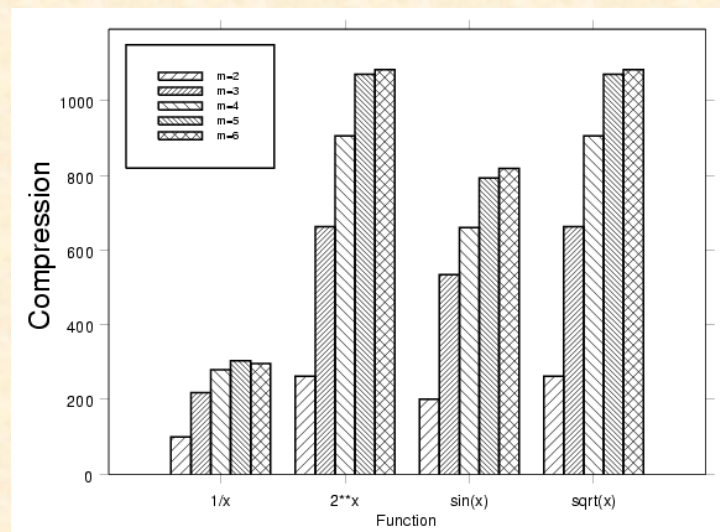
$f(x)$	Tables	Bit Partitions	Table Sizes	Total Memory
$1/x$	1	23	$2^{23} \cdot 23$	192,937,984
$1/x$	2	9, 7, 7	$2^{16} \cdot 25 + 2^{15} \cdot 9$	1,933,312
$1/x$	3	11, 3, 4, 5	$2^{14}(26 + 12) + 2^{15} \cdot 8$	884,736
$1/x$	4	11, 3, 3, 3, 3	$2^{14} \cdot 27 + 2^{13}(13 + 10 + 7)$	688,128
$1/x$	5	11, 3, 2, 2, 2, 3	$2^{14} \cdot 27 + 2^{12}(13 + 11 + 9) + 2^{13} \cdot 7$	634,880
$1/x$	6	11, 3, 1, 2, 2, 2, 2	$2^{14} \cdot 28 + 2^{11} \cdot 14 + 2^{12}(13 + 11 + 9 + 7)$	651,264
$\sqrt{x}$	1	23	$2^{23} \cdot 23$	192,937,984
$\sqrt{x}$	2	7, 7, 9	$2^{14} \cdot 25 + 2^{15} \cdot 10$	737,280
$\sqrt{x}$	3	8, 5, 5, 5	$2^{13} \cdot 26 + 2^{12}(12 + 7)$	290,816
$\sqrt{x}$	4	9, 3, 3, 4, 4	$2^{12} \cdot 27 + 2^{11} \cdot 14 + 2^{12}(11 + 7)$	212,992
$\sqrt{x}$	5	9, 3, 2, 3, 3, 3	$2^{12} \cdot 27 + 2^{10} \cdot 14 + 2^{11}(12 + 9 + 6)$	180,224
$\sqrt{x}$	6	9, 3, 2, 2, 2, 2, 3	$2^{12} \cdot 28 + 2^{10}(15 + 13 + 11 + 9) + 2^{11} \cdot 7$	178,176
$\sin(x)$	1	24	$2^{24} \cdot 24$	402,653,184
$\sin(x)$	2	8, 8, 8	$2^{16} \cdot 26 + 2^{15} \cdot 9$	1,998,848
$\sin(x)$	3	10, 4, 5, 5	$2^{14} \cdot (27 + 12 + 7)$	753,664
$\sin(x)$	4	10, 4, 3, 3, 4	$2^{14} \cdot 28 + 2^{12}(13 + 10) + 2^{13} \cdot 7$	610,304
$\sin(x)$	5	11, 2, 2, 3, 3, 3	$2^{13} \cdot 28 + 2^{12} \cdot 14 + 2^{13}(12 + 9 + 6)$	507,904
$\sin(x)$	6	11, 2, 2, 2, 2, 2, 3	$2^{13} \cdot 29 + 2^{12}(15 + 13 + 11 + 9) + 2^{13} \cdot 7$	491,520
$2^x$	1	24	$2^{24} \cdot 23$	385,875,968
$2^x$	2	8, 7, 9	$2^{15} \cdot 25 + 2^{16} \cdot 10$	1,474,560
$2^x$	3	9, 5, 5, 5	$2^{14} \cdot 26 + 2^{13}(12 + 7)$	581,632
$2^x$	4	10, 3, 3, 4, 4	$2^{13} \cdot 27 + 2^{12} \cdot 14 + 2^{13}(11 + 7)$	425,984
$2^x$	5	10, 3, 2, 3, 3, 3	$2^{13} \cdot 27 + 2^{11} \cdot 14 + 2^{12}(12 + 9 + 6)$	360,448
$2^x$	6	10, 3, 2, 2, 2, 2, 3	$2^{13} \cdot 28 + 2^{11}(15 + 13 + 11 + 9) + 2^{12} \cdot 7$	356,352

2008年3月15日  
星期六 (25)

## Compression Ratio (16 bit)

2008年3月15日  
星期六 (26)

## Compression ratio (24 bit)



2008年3月15日  
星期六 (27)

## Conclusion

- The STAM method uses the first two terms of the Taylor series to construct an approximation to an arbitrary differentiable function
- Considerable area reduction over simple table lookup
- High speed (only adders and memories required)

2008年3月15日  
星期六 (28)

## References

- J. E. Stine and M. J. Schulte, "The Symmetric Table Addition Method for Accurate Function Approximation," in Journal of VLSI Signal Processing, vol. 21, no. 2, pp. 167-177, June, 1999.
- Software for computing the tables is available on <http://www.cse.lehigh.edu/~caar/STAM.html>. A version suitable for gcc 3.3 is available from the CEG5010 website

2008年3月15日  
星期六 (29)

## Review Questions

- Compute the  $a_0$ ,  $a_1$  and  $a_2$  tables for a small example of your own construction. Download the software and check its correctness.
- Make sure you understand the four sources of errors described in the lecture and the constraint equation relating the parameters and the 2<sup>nd</sup> derivative of the function