

Project: Fixed Point Division

Mahendra Khinchi

Solution:

Mathematical Part:

0.1 Initial Setup

$$\begin{aligned}N &= 8 \\ \text{scale_factor} &= 2^8 = 256 \\ D &= 23 \quad (\text{denominator}) \\ \text{num} &= 15 \quad (\text{numerator}) \\ D_{\text{fixed}} &= \text{round}(23 \times 256) = 5888 \\ x_0 &= \frac{1}{17} \approx 0.0588235 \\ x_{\text{fixed}} &= \text{round}(0.0588235 \times 256) = 15\end{aligned}$$

0.2 Newton-Raphson Iterations

Formula:

$$\begin{aligned}x_{n+1} &= x_n \times (2 - D \times x_n) \\ x_{n+1}^{\text{fixed}} &= \left(x_n \times \left(2 \times \text{scale_factor} - \frac{x_n \times D_{\text{fixed}}}{2^N} \right) \right) \div 2^N\end{aligned}$$

Iterations:

$$\begin{aligned}x_1 &= (15 \times (512 - 345)) \div 256 = 9 \\ x_2 &= (9 \times (512 - 207)) \div 256 = 10 \\ x_3 &= (10 \times (512 - 230)) \div 256 = 11 \\ x_4 &= x_5 = (11 \times (512 - 253)) \div 256 = 11\end{aligned}$$

Final Calculation:

$$\begin{aligned}\text{result_fixed} &= \text{num} \times x_{\text{fixed}} = 15 \times 11 = 165 \\ \text{result_floating} &= \frac{165}{256} \approx 0.64453125\end{aligned}$$

1 MATLAB Implementation

```
1 N = 8;
2 scale_factor = 2^N;
3 D = 23; % Denominator
4 num = 15; % Numerator
5
6 % Convert to fixed-point
7 D_fixed = round(D * scale_factor); % Convert denominator to fixed
    point
8 x0 = 1 / 17;
9 x_fixed = round(x0 * scale_factor); % Convert initial guess 1/D to
    fixed point
10
11 % Newton-Raphson Iteration in Fixed-Point using bit shifts
12 for i = 1:5
13     x_fixed = bitshift(x_fixed * (2 * scale_factor - bitshift(
        x_fixed * D_fixed, -N)), -N);
14 end
15
16 % Compute 15/23 in fixed point using bit shifts
17 result_fixed = num * x_fixed;
18
19 % Convert to floating point for verification
20 result_floating = result_fixed / scale_factor;
21
22 % Display Results
23 disp(['Fixed-Point Result: ', num2str(result_fixed)]);
24 disp(['Floating-Point Result: ', num2str(result_floating)]);
```

2 Verilog Implementation

```
1 module fixed_point_division;
2     parameter N = 8;
3     parameter SCALE = 1 << N;
4     parameter D = 23;
5     parameter NUM = 15;
6
7     reg [16:0] D_fixed;
8     reg [16:0] x_fixed;
9     reg [16:0] temp;
10    integer i;
11
12    initial begin
13        D_fixed = D * SCALE; // Fixed-point representation of D
14        x_fixed = (SCALE / 17); // Initial approximation of 1/D
15
16        for (i = 0; i < 5; i = i + 1) begin
17            temp = x_fixed * (2 * SCALE - ((x_fixed * D_fixed) >> N
                ));
18            x_fixed = temp >> N;
19        end
20
21        // Multiply by NUM and compute the final result
```

```
22         temp = NUM * x_fixed;
23
24         // Display results
25         $display("D_fixed: %d", D_fixed);
26         $display("Final x_fixed: %d", x_fixed);
27         $display("Result Fixed: %d", temp);
28         $display("Result Floating: %f", temp * 1.0 / SCALE);
29
30         $stop;
31     end
32 endmodule
```