Bài tập Introduction

1. **Floating point.** Write 84.175, −528.685, 0.000924138, and −362005 in floating-point form, rounded to 5S (5 significant digits).
2. Write −76.437125, 60100, and −0.00001 in floating-point form, rounded to 4S.
3. **Small differences of large numbers** may be particularly strongly affected by rounding errors. Illustrate this by computing $0.81534/(35 \cdot 724 - 35.596)$ as given with 5S, then rounding stepwise to 4S, 3S, and 2S, where "stepwise" means round the rounded numbers, not the given ones.
4. **Order of terms**, in adding with a fixed number of digits, will generally affect the sum. Give an example. Find empirically a rule for the best order.

5. Rounding and adding. Let $a_1, \cdots, a_n$ be numbers with a_j correctly rounded to S_j digits. In calculating the sum $a_1 + \cdots + a_n$, retaining $S = \min S_j$ significant digits, is it essential that we first add and then round the result or that we first round each number to S significant digits and then add?

6. Nested form. Evaluate

$$\begin{aligned} f(x) &= x^3 - 7.5x^2 + 11.2x + 2.8 \\ &= ((x - 7.5)x + 11.2)x + 2.8 \end{aligned}$$

at $x = 3.94$ using 3S arithmetic and rounding, in both of the given forms. The latter, called the *nested form*, is usually preferable since it minimizes the number of operations and thus the effect of rounding.

7. Quadratic equation. Solve $x^2 - 30x + 1 = 0$ by (4) and by (5), using 6S in the computation. Compare and comment.

8. Solve $x^2 - 40x + 2 = 0$, using 4S-computation.

9. Do the computations in Prob. 7 with 4S and 2S.

10. Instability. For small $|a|$ the equation $(x - k)^2 = a$ has nearly a double root. Why do these roots show instability?

11. Theorems on errors. Prove Theorem 1(a) for addition.

12. Overflow and underflow can sometimes be avoided by simple changes in a formula. Explain this in terms of $\sqrt{x^2 + y^2} = x\sqrt{1 + (y/x)^2}$ with $x^2 \geqq y^2$ and x so large that x^2 would cause overflow. Invent examples of your own.

13. Division. Prove Theorem 1(b) for division.

14. Loss of digits. Square root. Compute $\sqrt{x^2 + 4} - 2$ with 6S arithmetic for $x = 0.001$ **(a)** as given and **(b)** from $x^2/(\sqrt{x^2 + 4} + 2)$ (derive!).

15. Logarithm. Compute $\ln a - \ln b$ with 6S arithmetic for $a = 4.00000$ and $b = 3.99900$ **(a)** as given and **(b)** from $\ln(a/b)$.

16. **Cosine.** Compute $1 - \cos x$ with 6S arithmetic for $x = 0.02$ **(a)** as given and **(b)** by $2 \sin^2 \frac{1}{2}x$ (derive!).

17. Discuss the numeric use of (12) in App. A3.1 for $\cos v - \cos u$ when $u \approx v$.

18. **Quotient near 0/0.** **(a)** Compute $(1 - \cos x)/\sin x$ with 6S arithmetic for $x = 0.005$. **(b)** Looking at Prob. 16, find a much better formula.

19. **Exponential function.** Calculate $1/e = 0.367879$ (6S) from the partial sums of 5–10 terms of the Maclaurin series **(a)** of e^{-x} with $x = 1$, **(b)** of e^x with $x = 1$ and then taking the reciprocal. Which is more accurate?

20. Compute e^{-10} with 6S arithmetic in two ways (as in Prob. 19).

21. **Binary conversion.** Show that

$$\begin{aligned} 23 &= 20 \cdot 10^1 + 3 \cdot 10^0 = 16 + 4 + 2 + 1 \\ &= 2^4 + 2^2 + 2^1 + 2^0 = (1 \quad 0 \quad 1 \quad 1 \quad 1.)_2 \end{aligned}$$

can be obtained by the division algorithm

$$\begin{array}{lll} 2\,\underline{\lfloor 23} & \text{Remainder} & 1 = c_0 \\ 2\,\underline{\lfloor 11} & & 1 = c_1 \\ 2\,\underline{\lfloor 5\ } & & 1 = c_2 \\ 2\,\underline{\lfloor 2\ } & & 0 = c_3 \\ \quad 0 & & 1 = c_4 \end{array}$$

22. Convert $(0.59375)_{10}$ to $(0.10011)_2$ by successive multiplication by 2 and dropping (removing) the integer parts, which give the binary digits $c_1, c_2, \cdots$:

$$
\begin{aligned}
& 0\ .59375 \cdot 2 \\
c_1 &= \boxed{1}\ .1875 \cdot 2 \\
c_2 &= \boxed{0}\ .375 \cdot 2 \\
c_3 &= \boxed{0}\ .75 \cdot 2 \\
c_4 &= \boxed{1}\ .5 \cdot 2 \\
c_5 &= \boxed{1}\ .0
\end{aligned}
$$

23. Show that 0.1 is not a binary machine number.

24. Prove that any binary machine number has a finite decimal representation. Is the converse true?

25. CAS EXPERIMENT. Approximations. Obtain $x = 0.1 = \frac{3}{2} \sum_{m=1}^{\infty} 2^{-4m}$ from Prob. 23. Which machine number (partial sum) S_n will first have the value 0.1 to 30 decimal digits?

26. CAS EXPERIMENT. Integration from Calculus. Integrating by parts, show that $I_n = \int_0^1 e^x x^n \, dx = e - nI_{n-1}$, $I_0 = e - 1$. **(a)** Compute I_n, $n = 0, \cdots,$ using 4S arithmetic, obtaining $I_8 = -3.906$. Why is this nonsense? Why is the error so large?

(b) Experiment in (a) with the number of digits $k > 4$. As you increase k, will the first negative value $n = N$ occur earlier or later? Find an empirical formula for $N = N(k)$.

27. Backward Recursion. In Prob. 26. Using $e^x < e$ $(0 < x < 1)$, conclude that $|I_n| \leqq e/(n+1) \to 0$ as $n \to \infty$. Solve the iteration formula for $I_{n-1} = (e - I_n)/n$, start from $I_{15} \approx 0$ and compute 4S values of $I_{14}, I_{13}, \cdots, I_1$.

28. Harmonic series. $1 + \frac{1}{2} + \frac{1}{3} + \cdots$ diverges. Is the same true for the corresponding series of computer numbers?

29. Approximations of $\pi = 3.14159265358979$ $\cdots$ are $22/7$ and $355/113$. Determine the corresponding errors and relative errors to 3 significant digits.

30. Compute π by Machin's approximation $16 \arctan\left(\frac{1}{5}\right) - 4 \arctan\left(\frac{1}{239}\right)$ to 10S (which are correct). [In 1986, D. H. Bailey (NASA Ames Research Center, Moffett Field, CA 94035) computed almost 30 million decimals of π on a CRAY-2 in less than 30 hrs. The race for more and more decimals is continuing. See the Internet under pi.]