

Foundations of Information Systems

Winter Semester 2023–24, Exercise 5

For discussion on Wednesday, November 22, 2023

1. Identify values of x for which there is a substantial growth of relative error when the formula is evaluated in floating point arithmetic (due to subtraction of almost equal numbers). Then suggest an alternate formula that improves accuracy for the problematic range of x .

(a) $\frac{1 - (1 - x)^3}{x}$

(b) $\frac{1 - \sqrt{1 - x^2}}{x}$

(c) $\frac{1 - \sec x}{\tan^2 x}$

Hint: Recall that $\sec x = (\cos x)^{-1}$; use the well-known trigonometric identity $\sec^2 x = \tan^2 x + 1$.

2. Describe an algorithm which takes an string containing upper and lower case characters encoded in ASCII and outputs the same string in all lower case.

For simplicity, assume that your input string does not contain any symbols other than characters of the Latin alphabet. You may describe your algorithm in pseudo-code or in Python.

3. How many characters can we encode with a mixed 8/16 bit representation if we insist that an 8 bit character always has 0 in the first bit, and a 16 bit character always has 1 in the first bit of each of its bytes?
4. List at least three advantages of the UTF-8 encoding over the UTF-16 encoding of Unicode characters.