

2002 Paper 10 Question 11

Numerical Analysis I

- (a) Consider a version of the Brown model in which the significand of a floating-point number is represented as $d_0.d_1d_2 \dots d_{p-1}$. Explain the parameters β , p , e_{\max} , e_{\min} of the model. [3 marks]
- (b) Describe the layout of bits in IEEE single precision and give the values of the above four parameters. [5 marks]
- (c) IBM System/370 single precision uses the same total number of bits, and a similar method for storing negative exponents. However, there are 7 bits for the exponent, and all bit patterns represent numbers. Given $\beta = 16$, deduce the values of the remaining three parameters for this floating-point implementation. [5 marks]
- (d) If $\beta = 10$, $p = 3$ how should 6.789, 6.785, 6.755 be rounded using the “round to even” method? [3 marks]
- (e) Now consider $\beta = 2$, $p = 8$ on a machine with just one guard digit. How should the following be rounded using “round to even”?

011010110
101110101
110100011
011111111

[4 marks]