Example 2-bit exp. & 2-bit stored significand

**FIGURE 2.9** Floating-point numbers near zero

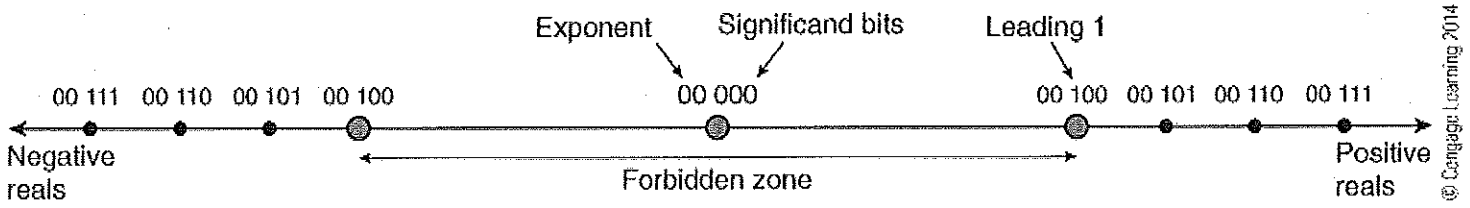





**FIGURE 2.10** Flowchart for floating-point addition and subtraction

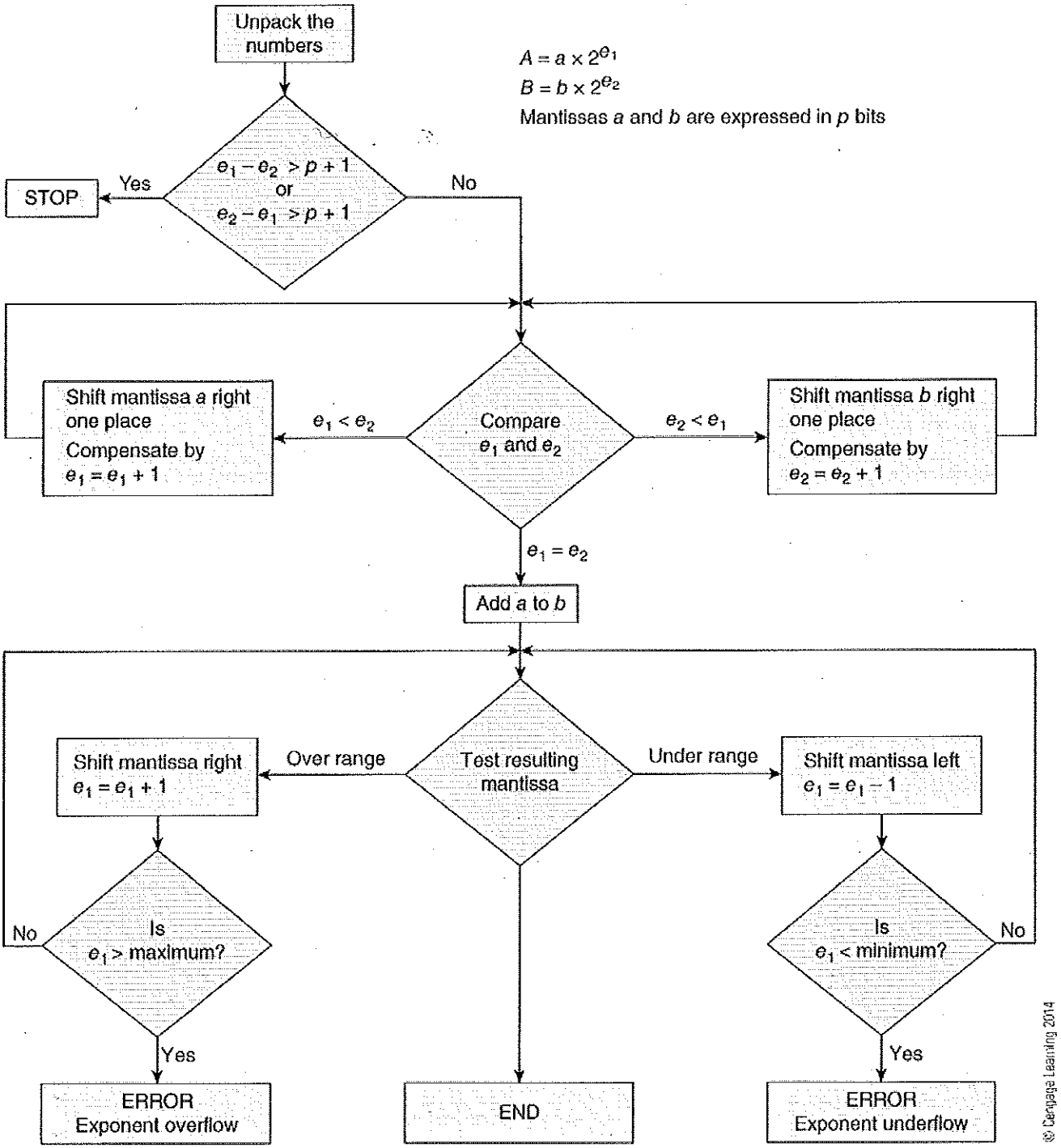
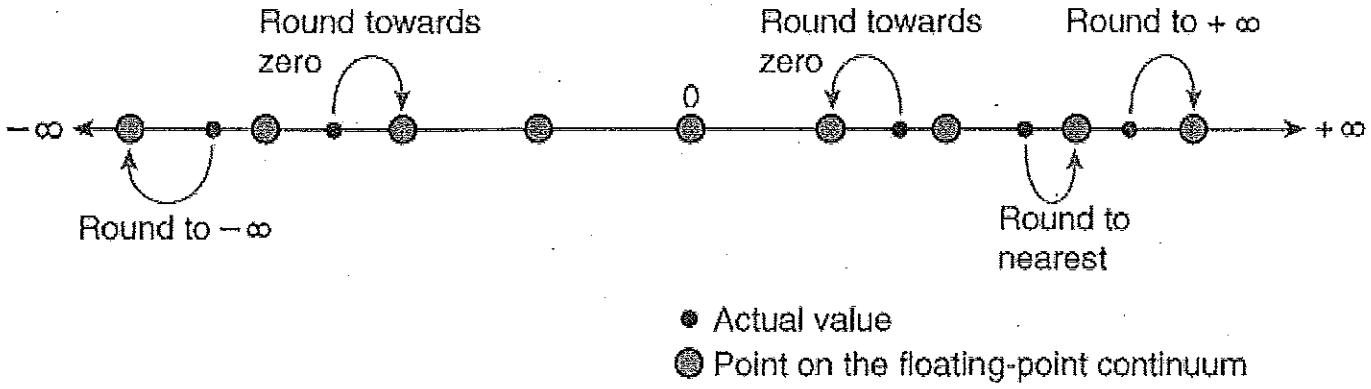


**FIGURE 2.11** Rounding mechanisms






## Other Features of IEEE Floating-Point Numbers

- The IEEE standard specifies that the default rounding technique should be towards the nearest value. If the number is equally distant from two nearest representable values, the value whose least-significant bit is zero is taken; that is, the rounding is towards the even value. The standard requires that three other rounding modes must be available (round toward zero, and round towards positive or negative infinity).
- The four comparisons specified by the IEEE format are equal to, less than, greater than, and *unordered*. The latter case, unordered, arises when one of the operands is a NaN.
- The IEEE format specifies five *exceptions*. An exception or interrupt is a request for attention that forces the computer to take a different course of action. The exceptions defined by the IEEE format are

  **Invalid Operation**—This exception arises when the programmer attempts something illegal such as an operation on a NaN or an addition or subtraction with infinity, or an attempt to calculate the square root of a negative number.

  **Division by Zero**—An exception is raised whenever an attempt is made to divide a number by zero (because the result is infinity).

  **Overflow**—Overflow exception occurs when the result is larger than the largest value that can be stored. Overflow can be dealt with by terminating the calculation or by using saturation arithmetic (holding the value at its maximum). Overflow in floating-point arithmetic is not the same as overflow in signed integer arithmetic.

  **Underflow**—Underflow exception occurs when the destination operand is smaller than the smallest value that can be stored; that is, the result is smaller than $2^{E_{min}}$. Underflow can be handled by setting the smallest number to zero or by representing the number as an unnormalized value less than $2^{E_{min}}$.

  **Inexact**—An inexact exception occurs when a round-off error is committed by an operation.

- NaNs *propagate* when used in expressions; that is, if part of an expression is a NaN, the result is a NaN.