First name: Monika
Last name: Wunderlich
Date: 4.12.03
Homeworknumber: 1
Homework Title: 1.17

Problem description:
Let x be a given nonzero floating-point number in a normalized system, and let y be an adjacent floating-point number, also nonzero.

(a) What is the minimum possible spacing between x and y?

(b) What is the maximum possible spacing between x and y?

Problem solution:
In a normalized system the floating-point number x is represented by the mantissa $d_0 d_1 \ldots d_{p-1}$ where d_0 is nonzero and the exponent $E \in [L, U]$.

$$x = \left(d_0 + \frac{d_1}{\beta} + \frac{d_2}{\beta^2} + \ldots + \frac{d_{p-1}}{\beta^{p-1}} \right) \cdot \beta^E$$

There are two possible cases for the lower adjacent floating-point number y:

$$y = \begin{cases} \left(d_0 + \frac{d_1}{\beta} + \frac{d_2}{\beta^2} + \ldots + \frac{d_{p-2}}{\beta^{p-2}} + \frac{d_{p-1}-1}{\beta^{p-1}} \right) \cdot \beta^E & \text{if } d_{p-1} > 0 \\ \left(d_0 + \frac{d_1}{\beta} + \ldots + \frac{d_i - 1}{\beta^i} + \frac{\beta-1}{\beta^{i+1}} + \ldots + \frac{\beta-1}{\beta^{p-1}} \right) \cdot \beta^E & \text{if } \forall j > i : d_j = 0 \\ & \text{and } d_i > 0,\ i > 0 \\ \left((\beta - 1) + \frac{\beta-1}{\beta} + \ldots + \frac{\beta-1}{\beta^{p-1}} \right) \cdot \beta^{E-1} & \text{if } \forall j > 0 : d_j = 0 \\ & \text{and } d_0 = 1 \end{cases}$$

Therefore we have three cases for calculation of the difference $\Delta = x - y$:

case: $d_{p-1} > 0$**:**

$$\begin{aligned} \Delta &= x - y \\ &= \left(d_0 + \frac{d_1}{\beta} + \ldots + \frac{d_{p-1}}{\beta^{p-1}} \right) \cdot \beta^E - \left(d_0 + \frac{d_1}{\beta} + \ldots + \frac{d_{p-2}}{\beta^{p-2}} + \frac{d_{p-1} - 1}{\beta^{p-1}} \right) \cdot \beta^E \\ &= \left(\frac{d_{p-1}}{\beta^{p-1}} - \frac{d_{p-1} - 1}{\beta^{p-1}} \right) \cdot \beta^E \\ &= \beta^{E-(p-1)} \end{aligned}$$

case: $\forall j > i : d_j = 0$ **and** $d_i > 0$ **and** $i > 0$**:**

$$\begin{aligned}
\Delta &= x - y \\
&= \left(d_0 + \frac{d_1}{\beta} + \ldots + \frac{d_i}{\beta^i} + \frac{0}{\beta^{i+1}} + \ldots + \frac{0}{\beta^{p-1}}\right) \cdot \beta^E \\
&\quad - \left(d_0 + \frac{d_1}{\beta} + \ldots + \frac{d_i - 1}{\beta^i} + \frac{\beta - 1}{\beta^{i+1}} + \ldots + \frac{\beta - 1}{\beta^{p-1}}\right) \cdot \beta^E \\
&= \left(\frac{1}{\beta^i} - \frac{\beta - 1}{\beta^{i+1}} - \frac{\beta - 1}{\beta^{i+2}} - \ldots - \frac{\beta - 1}{\beta^{p-1}}\right) \cdot \beta^E \\
&= \beta^{E-(p-1)}
\end{aligned}$$

case: $\forall j > 0 : d_j = 0$ **and** $d_0 = 1$**:**

$$\begin{aligned}
\Delta &= x - y \\
&= 1 \cdot \beta^E - \left((\beta - 1) + \frac{\beta - 1}{\beta} + \ldots + \frac{\beta - 1}{\beta^{p-1}}\right) \cdot \beta^{E-1} \\
&= \beta^{E-1-(p-1)} = \beta^{E-p}
\end{aligned}$$

We see that for the same x, Δ gets its minimum in the third case and its maximum in the first or second case.

Results:

(a) Thus $E \in [L, U]$ we get the minimum for Δ in the third case if E is minimal:

$$\Delta_{min} = \beta^{L-p}.$$

(b) The maximum of Δ is reached in the first or second case if E is maximal:

$$\Delta_{max} = \beta^{U-p+1}.$$