שאלה 1

בהינתן הפונקציה הבאה:

$$f(x_1,\dots,x_n) = \prod_{i=1}^{n} x_i^{\,i}$$

1. פתחו ביטוי עבור השגיאה המוחלטת והיחסית של f בהינתן השגיאות המוחלטות $\{\Delta x_i\}_{i=1}^{n}$.

1.1. $f(x_1,\dots,x_n) = \prod_{i=1}^{n} x_i^{\,i}$

$$\Delta\tilde{f} = \sum_{i=1}^{n} i\, x_i^{\,i-1}\cdot\Delta\tilde{x}_i\cdot\prod_{\substack{i=1\\ i\neq j}}^{n} x_j^{\,j}$$

$$\delta\tilde{f} = \frac{\Delta\tilde{f}}{f} = \frac{\sum_{i=1}^{n} i\, x_i^{\,i-1}\cdot\Delta\tilde{x}_i\cdot\prod_{\substack{i=1\\ i\neq j}}^{n} x_j^{\,j}}{\prod_{i=1}^{n} x_i^{\,i}} = \sum_{i=1}^{n}\frac{i\,\Delta\tilde{x}_i}{x_i} = \sum_{i=1}^{n} i\,\delta\tilde{x}_i$$

2. חשבו את ערך הפונקציה עבור $x_i = i,\ n = 5$.

2. $f(1,2,3,4,5) = 1^1\cdot 2^2\cdot 3^3\cdot 4^4\cdot 5^5 = 864\cdot 10^5$

3. חשבו את קירוב הפונקציה $\hat{f}$ עבור אותו n כאשר $\hat{x}_i = x_i + \frac{(-1)^{i+1}}{i^2}$.

3. $f(\tilde{1},\tilde{2},\tilde{3},\tilde{4},\tilde{5}) = \left(1+\frac{1}{1}\right)^1\cdot\left(2-\frac{1}{4}\right)^2\cdot\left(3+\frac{1}{9}\right)^3\cdot\left(4-\frac{1}{16}\right)^4\cdot\left(5+\frac{1}{25}\right)^5 = 144174219.2$

4. חשבו את השגיאה המוחלטת והיחסית עבור $\hat{f}$.

4. $\Delta\tilde{f} = |864\cdot 10^5 - 144174219.2| = 57774219.2$

$\delta\tilde{f} = \frac{\Delta\tilde{f}}{f} = 0.67 = 67\%$

5. חשבו את החסמים לשגיאות של $\hat{f}$ בעזרת הביטויים אותם פיתחתם בסעיף 1.

5. $\Delta\tilde{f} = \sum_{i=1} i\, x_i^{\,i-1}\Delta\tilde{x}_i\cdot\prod_{\substack{i=1\\ i\neq j}}^{5} x_j^{\,j} = 1\cdot 1^0\cdot 1\cdot\left(2^2\cdot 3^3\cdot 4^4\cdot 5^5\right) + 2\cdot 2^1\cdot\frac{1}{4}\cdot\left(1^1\cdot 3^3\cdot 4^4\cdot 5^5\right)$

$+ 3\cdot 3^2\cdot\frac{1}{9}\left(1^1\cdot 2^2\cdot 4^4\cdot 5^5\right) + 4\cdot 4^3\cdot\frac{1}{16}\left(1^1\cdot 2^2\cdot 3^3\cdot 5^5\right)$

$+ 5\cdot 5^4\cdot\frac{1}{25}\left(1^1\cdot 2^2\cdot 3^3\cdot 4^4\right) = 864\cdot 10^5 + 216\cdot 10^5$

$+ 96\cdot 10^5 + 54\cdot 10^5 + 3456000 = 126456000$

$\delta\tilde{f} = \sum_{i=1}^{5} i\,\delta\tilde{x}_i = 1\cdot\frac{1}{1} + 2\cdot\frac{\frac{1}{4}}{2} + 3\cdot\frac{\frac{1}{9}}{3} + 4\cdot\frac{\frac{1}{16}}{4} + 5\cdot\frac{\frac{1}{25}}{5} = 1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \frac{1}{25} = 1.464$

$= 146.4\%$

שאלה 2

סטודנט בנה תוכנית המדפיסה את כל הפרמוטציות האפשריות של המספרים 1 עד 11. כדי לוודא שהתוכנית הדפיסה את כל הפרמוטציות הוסיף הסטודנט מונה שערכו ההתחלתי 0 ובעבור כל הדפסה גדל ב-1.
הסטודנט הגדיר את המונה כמשתנה נקודה צפה מסוג Single כפי שמוגדר בIEEE 754.

1. בהנחה שהתוכנית מדפיסה את כל הפרמוטציות(11!), מה יהיה ערכו של המונה בסיום התוכנית? רשמו ייצוג בינארי של המספר(מנטיסה, חזקה וסיבית הסימן).

1. 2^{24} would be the counter final value, since we increment the counter from this point onwards, there would be loss of significance $2^{24}+1=2^{24}$.

$S=0$ $\quad exp=24+127=151_{(10)}=10010111_{(2)}$ $\quad fraction=0\ldots0$

2. רשמו את השגיאה המוחלטת והיחסית בערך המונה.

2. $X=11!=39916800$, $\tilde{X}=2^{24}$

$\Delta\tilde{X}=|39916$

$|11!-2^{24}|=|39916800-16777216|=23139584$

$\delta\tilde{X}=\frac{\Delta\tilde{X}}{|X|}=\frac{23139584}{39916800}=0.5797=57.97\%$

3. האם ניתן לתקן את הבעיה בעזרת שימוש משתנה נקודה צפה נוסף מסוג Single? הסבירו את תשובתכם.

3. Yes, we could use another variable as a counter of 2^{24}'s

```
X++  ⟶   if x = 2^24 - 1 :
              y++
              x = 0
          else:
              X++
```

where y is our added counter.

שאלה 3

פתחו ביטויים המתארים את השגיאות היחסיות והמוחלטות עבור הפונקציות הבאות:

1. $f_1(x,y) = y^{e^x}$

$$\Delta\tilde{f}_1 = \left|\frac{\partial f_1}{\partial x}\right|\cdot\Delta\tilde{x} + \left|\frac{\partial f_1}{\partial y}\right|\cdot\Delta\tilde{y} = \left|e^x\cdot\ln y\cdot y^{e^x}\right|\cdot\Delta\tilde{x} + \left|e^x y^{e^x-1}\right|\Delta\tilde{y}$$

$$\delta\tilde{f}_1 = \frac{\Delta\tilde{f}_1}{|f_1|} = \frac{\left|e^x\ln y\cdot y^{e^x}\right|\cdot\Delta\tilde{x} + \left|e^x y^{e^x-1}\right|\Delta\tilde{y}}{\left|y^{e^x}\right|} = \left|\frac{e^x\ln y\cdot y^{e^x}}{y^{e^x}}\right|\cdot\Delta\tilde{x} + \left|\frac{e^x y^{e^x-1}}{y^{e^x}}\right|\cdot\Delta\tilde{y}$$

$$= e^x\ln y\cdot\Delta\tilde{x} + e^x\cdot\delta\tilde{y}$$

2. $f_2(x,y,z) = x^2 2^y + ln(z)$

$$\Delta\tilde{f}_2 = \left|\frac{\partial f_2}{\partial x}\right|\cdot\Delta\tilde{x} + \left|\frac{\partial f_2}{\partial y}\right|\cdot\Delta\tilde{y} + \left|\frac{\partial f_2}{\partial z}\right|\cdot\Delta\tilde{z} = \left|2^{y+1}\cdot x\right|\Delta\tilde{x} + \left|\ln(2)x^2\cdot 2^y\right|\cdot\Delta\tilde{y} + \left|\frac{1}{z}\right|\cdot\Delta\tilde{z}$$

$$= \left|2^{y+1}\cdot x\right|\cdot\Delta\tilde{x} + \left|\ln(2)x^2\cdot 2^y\right|\Delta\tilde{y} + \delta\tilde{z}$$

$$\delta\tilde{f}_2 = \frac{\Delta\tilde{f}_2}{|f_2|} = \frac{\left|2^{y+1}\cdot x\right|\cdot\Delta\tilde{x} + \left|\ln(2)\cdot x^2\cdot 2^y\right|\cdot\Delta\tilde{y} + \delta\tilde{z}}{\left|x^2 2^y + \ln(z)\right|}$$

3. $f_3(x,y,z) = \frac{7(x^4+y^3z^2)^5}{x-z^y}$

$$\Delta\tilde{f}_3 = \left|\frac{\partial f_3}{\partial x}\right|\cdot\Delta\tilde{x} + \left|\frac{\partial f_3}{\partial y}\right|\cdot\Delta\tilde{y} + \left|\frac{\partial f_3}{\partial z}\right|\cdot\Delta\tilde{z}$$

$$= \left|\frac{35\cdot 4x^3(x^4+y^3z^2)^4(x-z^y) - 7(x^4+y^3z^2)^5}{(x-z^y)^2}\right|\cdot\Delta\tilde{x}$$

$$+ \left|\frac{35\cdot 3y^2z^2(x^4+y^3z^2)^4}{x-z^y}\right|\cdot\Delta\tilde{y}$$

$$+ \left|\frac{35\cdot 2zy^3(x^4+y^3z^2)^4(x-z^y) - (-yz^{y-1})(7(x^4+y^3z^2)^5)}{(x-z^y)^2}\right|\cdot\Delta\tilde{z}$$

$$= \left|\frac{7(x^4+y^3z^2)^4(20x^3(x-z^y) - x^4 - y^3z^2)}{(x-z^y)^2}\right|\cdot\Delta\tilde{x}$$

$$+ \left|\frac{105y^2z^2(x^4+y^3z^2)^4}{x-z^y}\right|\cdot\Delta\tilde{y}$$

$$+ \left|\frac{7(x^4+y^3z^2)^4\cdot(10zy^3(x-z^y) + yz^{y-1}(x^4+y^3z^2))}{(x-z^y)^2}\right|\cdot\Delta\tilde{z}$$

$$\delta\tilde{f}_3 = \frac{\Delta\tilde{f}_3}{|f_3|} = \left|\frac{7(20x^3(x-z^y) - x^4 - y^3z^2)}{(x^4+y^3z^2)(x-z^y)}\right|\cdot\Delta\tilde{x} + \left|\frac{105y^2z^2}{x^4+y^3z^2}\right|\cdot\Delta\tilde{y} +$$

$$+ \left|\frac{7(10zy^3(x-z^y) + yz^{y-1}(x^4+y^3z^2))}{(x^4+y^3z^2)(x-z^y)}\right|\cdot\Delta\tilde{z}$$

$$= 7\cdot\left|\frac{20x^3}{x^4+y^3z^2} - \frac{1}{x-z^y}\right|\cdot\Delta\tilde{x} + \left|\frac{105y^2z^2}{x^4+y^3z^2}\right|\cdot\Delta\tilde{y} + 7\left|\frac{10zy^3}{x^4+y^3z^2} + \frac{yz^{y-1}}{x-z^y}\right|\cdot\Delta\tilde{z}$$

שאלה 4

בהינתן שיטת נקודה צפה חדשה בשם triple עם 96 סיביות:

- שדה הסימן - בגודל סיבית אחת.
- שדה הexponent - בגודל 20 סיביות.
- שדה הfraction - בגודל 75 סיביות.

השיטה עובדת בצורה דומה לשאר השיטות, כאשר exponent = 0 שמור למספרים לא מנורמלים, והexponent המקסימלי שמור לאינסוף וכו'.

1. חשבו את הbias של השיטה הנתונה.

1. $bias = \frac{2^{20}-2}{2} = 2^{19}-1 = 524287$

2. חשבו את המספרים הנמוכים והגבוהים ביותר במספרים המנורמלים והלא מנורמלים.

2. Normalized:

Lowest: $exp = 1 \rightarrow 1-524287 = -524286$

$fraction = 0\ldots0$

$X = 2^{-524286}$

Highest $exp = 1048574 \rightarrow 1048574 - 524287 = 524287$

$fraction = 1\ldots1$

$X = (2-2^{-75}) \cdot 2^{524287}$

Non-Normalized:

Lowest: $fraction = 0\ldots01$

$X = 2^{-75} \cdot 2^{-524286} = 2^{-524361}$

Highest: $fraction = 1\ldots1$

$X = (1-2^{-75}) \cdot 2^{-524286}$

3. ייצגו את שלושת המספרים הבאים או קירוביהם בעזרת שיטה זו:

- 624.5625

3. $624.5625_{(10)} = 1001110000.1001_{(2)} = 1.001110000 1001 \cdot 2^{9}$

$s=0 \quad exp = 9+524287 = 524296 = 1\underbrace{0\ldots0}_{\times 15}1000 \quad fraction = 0011100001001\underbrace{0\ldots0}_{\times 62}$

- 11.375×10^{10}

$11.375 \times 10^{10} \approx 11.375 \cdot 2^{33} = 1011.011 \cdot 2^{33} = 1.011011 \cdot 2^{36}$

$s=0,\ exp = 36+524287 = 1\underbrace{0\ldots0}_{\times 13}100011_,\ fraction = 011011\underbrace{0\ldots0}_{\times 69}$

- 6548×2^{-1994}

$6548 \times 2^{-1994} = 1100110010100 \times 2^{-1994} = 1.1001100101 \times 2^{-1982}$

$s=0,\ exp = -1982+524287 = 522305 = 0\underbrace{1\ldots1}_{\times 8}0000100001\ \ fraction = 1001100101\underbrace{0\ldots0}_{\times 65}$

שאלה 5

במערכת IEEE עם single precision, מהו ה-n הגדול ביותר עבורו מתקיים:

- $2^n + n^2 \neq 2^n$

2.2. * $n=33$: $2^{33} + 33^2 \neq 2^{33}$

since $33^2 = 1089$ needs 11 bits to be represented, so $33 - 10 = 23 \leq 23$ bits of representation we have

- $2^n + 2^{-n} \neq 2^n$

* $n=11$

since $2^{11} + 2^{-11}$ uses all available representation bits.