

Sprawozdanie z Laboratorium

Obliczenia Naukowe - Lista 1

Wojciech Typer

12 października 2025

1 Zdanie 1

1.1 Epsilon maszynowy (*macheps*)

Epsilonem maszynowym *macheps* nazywamy najmniejszą liczbę dodatnią taką, że w arytmetyce zmiennoprzecinkowej zachodzi $1.0 + \text{macheps} > 1.0$. Jest to miara precyzji obliczeń, która określa odległość od liczby 1.0 do następnej reprezentowalnej liczby maszynowej. Im mniejszy epsilon, tym większa precyzja arytmetyki, co jest bezpośrednio związane z liczbą bitów przeznaczonych na mantysę w danym typie zmiennoprzecinkowym.

Poniżej przedstawiono porównanie wartości *macheps* uzyskanych iteracyjnie, wartości zwracanych przez funkcję `eps()` w Julii oraz wartości zdefiniowanych w pliku nagłówkowym `float.h` kompilatora C (GCC 13).

Tabela 1: Porównanie wartości epsilon maszynowego.

Typ danych	Wartość z <code>float.h</code> (GCC)	Wartość z <code>eps(T)</code> (Julia)	Wartość wyznaczona iteracyjnie
Float16	$9.7656e-4$	$9.77e-4$	$9.77e-4$
Float32	$1.192\,092\,90e-7$	$1.192\,092\,9e-7$	$1.192\,092\,9e-7$
Float64	$2.220\,446\,049\,250\,313\,1e-16$	$2.220\,446\,049\,250\,313e-16$	$2.220\,446\,049\,250\,313e-16$

Jak widać w tabeli 1, wartości uzyskane eksperymentalnie są zgodne z wartościami referencyjnymi.

1.2 Najmniejsza dodatnia liczba maszynowa (*eta*)

Liczba *eta* (η) to najmniejsza dodatnia wartość, jaką można reprezentować w danym standardzie zmiennoprzecinkowym. Wartość ta jest związana z liczbami subnormalnymi (denormalizowanymi), które pozwalają na płynne "wypełnienie" luki między zerem a najmniejszą dodatnią liczbą znormalizowaną.

- **Związek z MIN_{sub} :** Liczba *eta* jest tożsama z MIN_{sub} , czyli najmniejszą możliwą do reprezentowania dodatnią liczbą subnormalną. W języku Julia wartość tę można uzyskać za pomocą funkcji `nextfloat(T(0.0))`.
- **Związek z MIN_{nor} :** Funkcja `floatmin(T)` zwraca najmniejszą dodatnią liczbę **znormalizowaną**, znaną jako MIN_{nor} . Jest to wartość większa od *eta*.

Wartości *eta* wyznaczone iteracyjnie (poprzez dzielenie 1.0 przez 2 aż do uzyskania 0.0) są zgodne z wynikami funkcji `nextfloat(T(0.0))`. Porównanie tych wartości przedstawiono w tabeli 2.

Tabela 2: Porównanie wartości η (η).

Typ danych	Wartość z <code>nextfloat(T(0.0))</code>	Wartość wyznaczona iteracyjnie
Float16	$6.0e-8$	$6.0e-8$
Float32	$1.4e-45$	$1.0e-45$
Float64	$5.0e-324$	$5.0e-324$

1.3 Największa wartość skończona (MAX)

Liczba MAX to największa skończona wartość, jaką można zapisać w danym typie zmiennoprzecinkowym. Próba reprezentacji liczby większej niż MAX prowadzi do uzyskania wartości nieskończonej (Inf). Doświadczenie wyznaczenie tej wartości polega na iteracyjnym mnożeniu liczby przez 2, aż do momentu, gdy stanie się ona nieskończona, a następnie cofnięciu ostatniej operacji.

Tabela 3: Porównanie maksymalnych wartości zmiennoprzecinkowych.

Typ danych	Wartość z <code>float.h</code> (GCC)	Wartość wyznaczona iteracyjnie
Float16	$6.550\,40e4$	$6.55e4$
Float32	$3.402\,823\,47e38$	$3.402\,823\,5e38$
Float64	$1.797\,693\,134\,862\,315\,7e308$	$1.797\,693\,134\,862\,315\,7e308$

2 Zadanie 2

W tabeli poniżej znajdują się wartości epsilon maszynowego, obliczone metodą Kahana i te, zwrócone przez funkcję `eps()` w Julii.

Typ danych	Wartość z metody Kahana	Wartość z <code>eps(T)</code> (Julia)
Float16	$-9.77e-4$	$9.77e-4$
Float32	$1.192\,092\,9e-7$	$1.192\,092\,9e-7$
Float64	$-2.220\,446\,049\,250\,313e-16$	$2.220\,446\,049\,250\,313e-16$

3 Zadanie 3