



Einige Beispiele zur numerischen Berechnung von Arithmetik-Termen im

Computeralgebra System Maxima



1 Vorbemerkungen



Das CAS Maxima verfügt über die
Zahlenformate integer, float und bfloat
("bigfloat").

Die Ergebnisse von Rechnungen, die strikt im
Ganzzahlbereich stattfinden, sind stets korrekt.



Maxima kann mit Brüchen rechnen. Auch die
Ergebnisse von Bruchrechnungen
auf der Basis von ganzen Zahlen sind stets
korrekt.



$1/3+1/6;$

$$\frac{1}{2}$$

Enthält eine Zahl einen (Dezimal-) Punkt, so ist sie automatisch eine Gleitkommazahl.

Die Ergebnisse von Rechnungen mit Gleitkommazahlen sind in der Regel nur näherungsweise (also nicht strikt) korrekt. Jede Gleitkommazahl (float oder bfloat) wird im Folgenden als verseucht bezeichnet.

Ein Term kann in gemischter Form ganze Zahlen und Gleitkommazahlen enthalten; m.a.W.. er kann unverseuchte und verseuchte Zahlen enthalten.

Die Qualität des Ergebnisses ist nur so gut wie die des "schlechtesten" Teilterms; m.a.W., ist ein Teilterm verseucht, so ist der ganze Term verseucht.

Mit Hilfe des Gleitkommaformats bfloat kann die Qualität von Berechnungen erheblich (genauer: fast beliebig) gesteigert werden.

Einige Beispiele:

$3/7 + 5/13$;

$$\frac{74}{91}$$

$3.0/7 + 5/13$;

0.8131868131868132

$\text{bfloat}(3.0/7) + \text{bfloat}(5/13)$;

8.131868131868132b-1

$(3/7 + 5/13) - (3.0/7 + 5/13)$;

0.0

$\text{is}((3/7 + 5/13) = (3.0/7 + 5/13))$;

false

$\text{is}((3/7 + 5/13) = (\text{bfloat}(3/7) + \text{bfloat}(5/13)))$;

false

$\text{is}((3/7 + 5/13) = 74/91)$;

true

$1/2 + 1/3 + 1/6;$

1

$\text{is}(1/2 + 1/3 + 1/6 = 1) ;$

true

$\text{is}(1.0/2 + 1/3 + 1/6 = 1) ;$

false

2 Beispiel: Term-Auswertung

Wir betrachten die Terme

$z1 : x*x*x*x - 4*y*y*y*y - 4*y*y \ \$$

$z21 : (x^2)^2 - 4*(y^2)^2 - 4*y^2 \ \$$

$z22 : x^4 - 4*y^4 - 4*y^2 \ \$$

$z3 : \exp(4*\log(x)) - 4*\exp(4*\log(y)) - 4*\exp(2*\log(y)) \ \$$

Rein mathematisch betrachtet, sollten alle vier

Terme für konkrete (reelle) Werte

von x und y dieselben Werte liefern: $z1 = z21$

$= z22 = z3$

Bemerkung: Die Werte von $z21$ und $z22$

wurden separat behandelt, da

in manchen Programmiersprachen das

Quadrieren anders gehandhabt wird

als das allgemeine Potenzieren.

Wir betrachten nun die Werte

$x : 665857.0 \ \$$

$y : 470832.0 \ \$$

Die korrekten Werte der obigen Terme für diese

Werte von x und y sind:

$z1 = z21 = z22 = z3 = 1$

Nun zu einigen konkreten Auswertungsverfahren

2.1 Auswertung durch die gewöhnliche floating point Arithmetik von Maxima

```

x : 665857.0 $
y : 470832.0 $
z1 : x*x*x*x - 4*y*y*y*y - 4*y*y;
z21 : (x^2)^2 - 4*(y^2)^2 - 4*y^2;
z22 : x^4 - 4*y^4 - 4*y^2;
z3 : exp(4*log(x)) - 4* exp(4*log(y)) - 4* exp(2*log(y));
      1.1885568 * 107
      1.1885568 * 107
      1.1885568 * 107
      -4.914309120002441 * 108

```

2.2 Auswertung durch die Ganzzahl-Arithmetik von Maxima

```

xi : 665857 $
yi : 470832 $
z1 : xi*xi*xi*xi - 4*yi*yi*yi*yi - 4*yi*yi;
z21 : (xi^2)^2 - 4*(yi^2)^2 - 4*yi^2;
z22 : xi^4 - 4*yi^4 - 4*yi^2;
z3 : exp(4*log(xi)) - 4*exp(4*log(yi)) - 4*exp(2*log(yi));
      1
      1
      1
      1

```

2.3 Auswertung durch die Bigfloat-Arithmetik von Maxima (floating point precision: fpprec = 16 default value)


```
10 0.1 1.0b-1 false
    done
```

4 *Eine einfache Iteration*

In der folgenden Iteration sollte stets $x = 0,2$ sein.

```
simple_iteration_fp(n) :=
  (x : 0.2,
   for i:1 thru n do
     (x : 11*x-2,
      print(i, x)) ) ;
```

```
    simple_iteration_fp(n):=
  (x:0.2,for i thru n do (x:11*x-2,print(i,x)))
```

```
simple_iteration_fp(10);
```

```
1 0.20000000000000002
2 0.20000000000000002
3 0.20000000000000215
4 0.20000000000002364
5 0.2000000000026008
6 0.2000000000286084
7 0.2000000003146924
8 0.2000000034616169
9 0.2000000380777864
10 0.2000004188556508
```

```
    done
```

```
simple_iteration_bf(n) :=
  (fpprec : 16,
   x : bfloat(0.2),
   for i:1 thru n do
     (x : 11*x-2,
      print(i, " ", x)) ) ;
```

```
    simple_iteration_bf(n):=(fpprec:16,x:bfloat(0.2),for i
thru n do (x:11*x-2,print(i, ,x)))
```

```
simple_iteration_bf(20);
```

```
1 2.0000000000000001b-1
2 2.0000000000000014b-1
3 2.000000000000148b-1
4 2.000000000001626b-1
5 2.00000000001788b-1
6 2.000000000196683b-1
```

[illegible]

[illegible]

[illegible]

[illegible][illegible]

done

```
(r1 : bfloat( (1/10 + 2/10) + 3/10 ) ,  
r2 : bfloat( 1/10 + (2/10 + 3/10) ) ,  
print(r1) ,  
print(r2) ,  
print(r1-r2),  
is(r1=r2) );
```

6.0b-1

6.0b-1

0.0b0

true

5.2 Das Distributivgesetz



```
( r1 : 100 * (0.1 + 0.2),
  r2 : (100 * 0.1 + 100 * 0.2),
  print(r1),
  print(r2),
  print(r1-r2),
  is(r1=r2) );
```

30.0

30.0

3.552713678800501 * 10⁻¹⁵*false*

```
( r1 : bfloat(100 * (1/10 + 2/10) ),
  r2 : bfloat(100 * 1/10 + 100 * 2/10 ),
  print(r1),
  print(r2),
  print(r1-r2),
  is(r1=r2) );
```

3.0b1

3.0b1

0.0b0

true