

Zad 8.

niedziela, 19 marca 2023 14:58

Zadanie 8. Napisz ciało procedury o sygnaturze «`int32_t int2float(int32_t i)`» konwertującej wartość całkowitoliczbową do binarnej reprezentacji liczby typu «float». Wynik należy zaokrąglić do najbliższej liczby parzystej – w tym celu wyznaczyć wartość bitów guard, round i sticky. Do wyznaczenia pozycji wiodącej jedynek można użyć funkcji «`__builtin_clz`», tj. [instrukcji wbudowanej](#)¹ w kompilator gcc.

Wskazówka: Rozwiązanie jest łatwiejsze do zrozumienia, jeśli mantysę zapisze się na najbardziej znaczących bitach.

Built-in Function: `int __builtin_clz (unsigned int x)`

Returns the number of leading 0-bits in *x*, starting at the most significant bit position. If *x* is 0, the result is undefined.

```
30 int32_t int2float (int32_t n) {
31     if (n == 0)
32         return 0;
33
34     int32_t sign = n & 0x80000000;
35
36     int32_t leading = __builtin_clz(n); // # of leading 0's
37     uint32_t e = (32 - (leading + 1) + 127); // exponent + bias
38
39     if (e >= 128)
40         return sign | (0xFF << 23); // +/- inf if n too large
41
42     uint32_t m = n << (leading + 1); // mantisa shifted to the left (ignoring leading 1)
43
44     uint32_t g = (m >> 9) & 1; // guard bit
45     uint32_t r = (m >> 8) & 1; // round bit
46     uint32_t s = (m << 24) > 0; // sticky bit
47
48     m = m >> 9; // 23 bits of mantisa shifted right
49
50     // rounding rules
51     if (r == 1 & (s == 1 | g == 1)) // increment if > half or == half and r = 1
52         m = m + 1;
53     m = m & 0x7FFFFFFF;
54
55     return sign | (e << 23) | m;
56 }
```