Course No : CSE 306

Assignment on Floating Point Adder

Section : B2

Group No : 5

Student IDs : 1705095
1705096
1705097
1705102
1705104

Introduction:

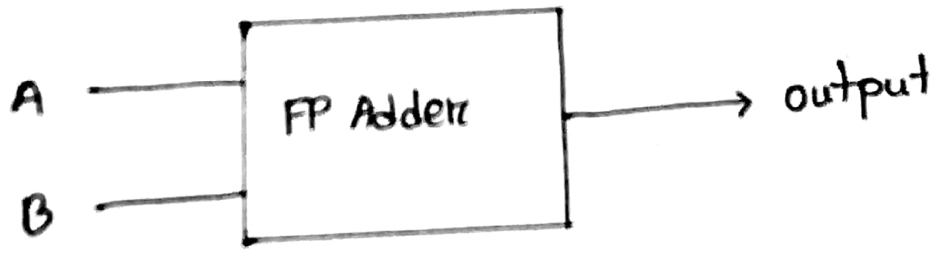
A floating point Adder is a combinational arithmatic logic unit which performs addition operation on signed floating point numbers. IEEE standardized several formats to represent floating point numbers, "half-precision", "single precision" etc. In this experiment, we simulate a floating point adder which takes two "half-precision" format floating point numbers as input and produce output in the same format.

Problem Specification:

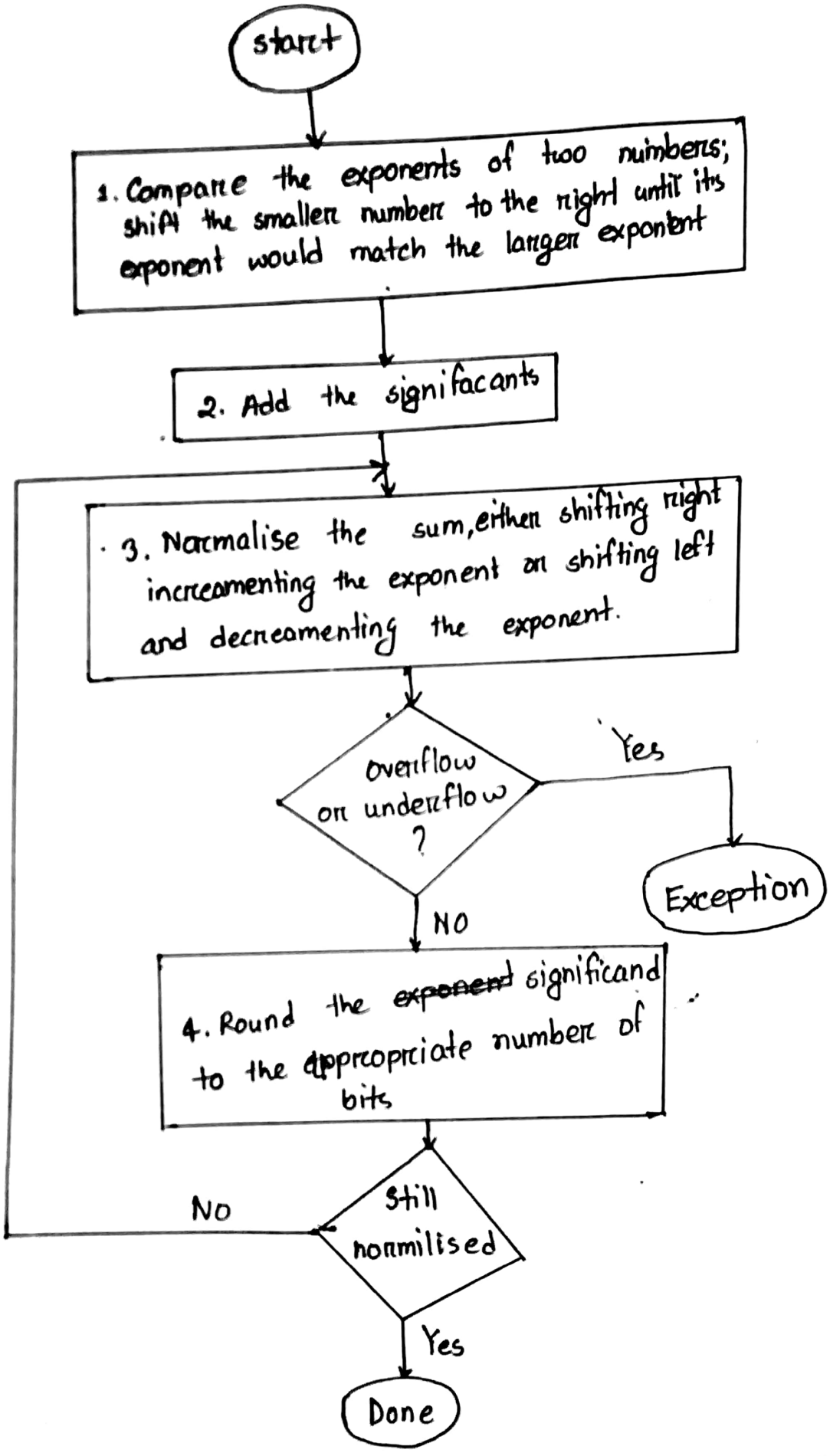
Here, we are required to design a floating point adder circuit which takes two floating points as inputs and provides their sum, another floating point as output. Each floating point will have to be 16 bits long with following representation

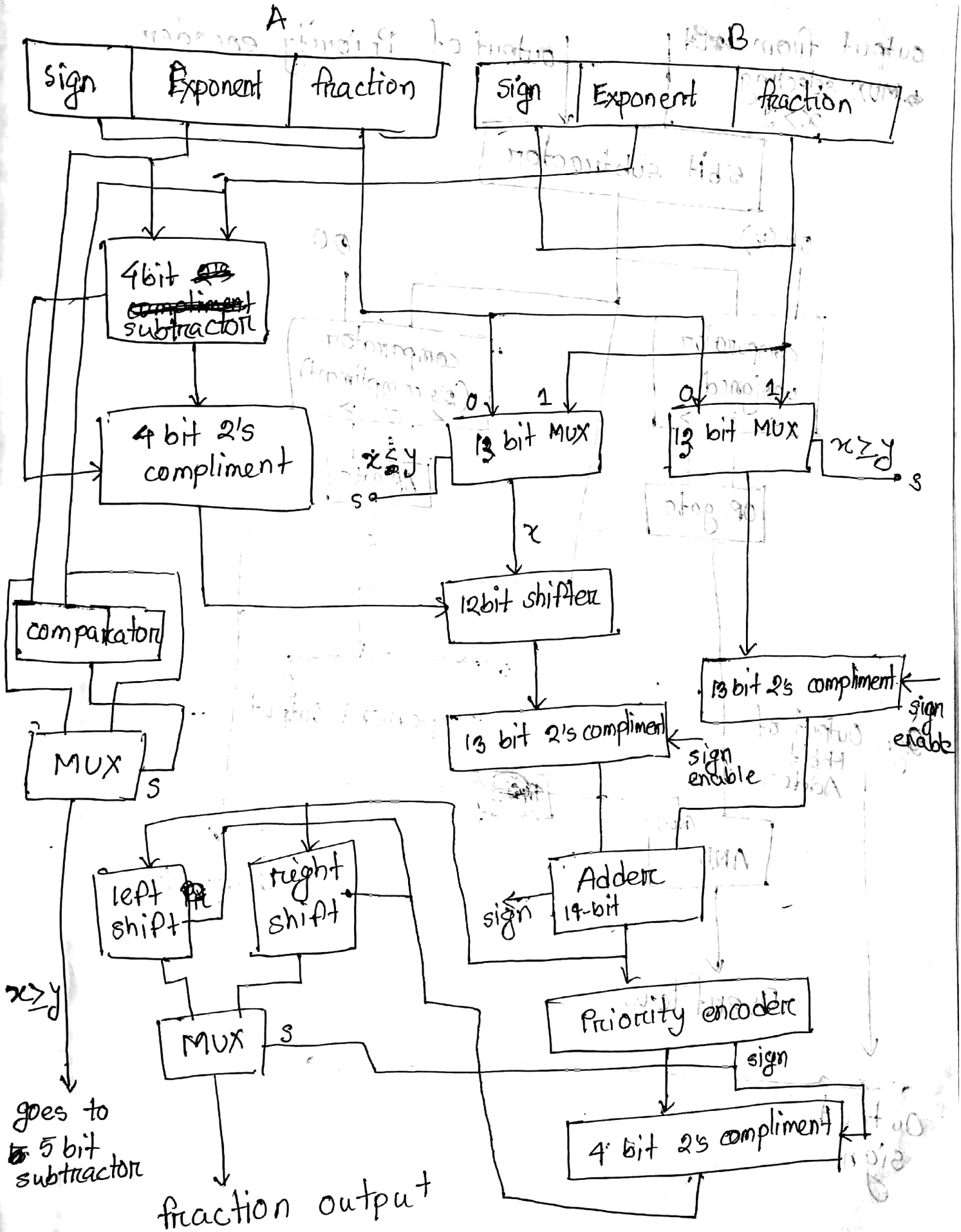
Sign	Exponent	Fraction
1bit	4bit	11 bit

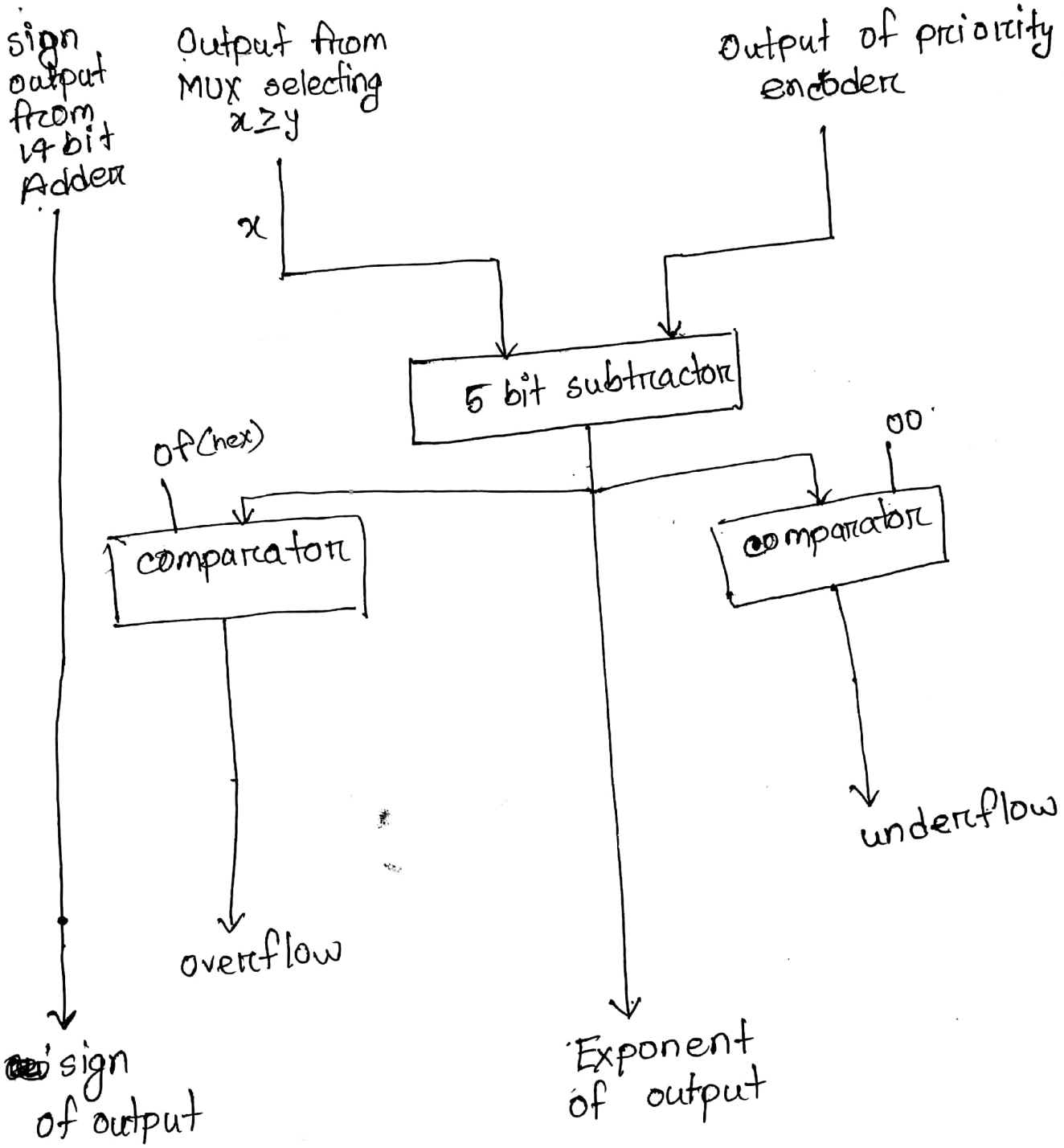
And the floating point adder,



Flowchart of the Algorithm:




A
B
sign
Exponent
fraction
sign
Exponent
fraction
4bit subtraction
4 bit 2's compliment
0
1
13 bit MUX
x ≤ y
s
0
1
13 bit MUX
x ≥ y
s
x
comparator
12bit shifter
MUX
s
13 bit 2's compliment
sign enable
13 bit 2's compliment
sign enable
left shift
right shift
Adder 14-bit
sign
x > y
MUX
s
Priority encoder
sign
4 bit 2's compliment
goes to 5 bit subtractor
fraction output


sign output from 14 bit Adder
Output from MUX selecting x≥y
Output of priority encoder
x
5 bit subtractor
of(hex)
00
comparator
comparator
overflow
underflow
sign of output
Exponent of output

Truth table for Priority encoder:

I_{12}	I_{11}	I_{10}	I_9	I_8	I_7	I_6	I_5	I_4	I_3	I_2	I_1	I_0	O_3	O_2	O_1	O_0
1	x	x	x	x	x	x	x	x	x	x	x	x	1	1	1	1
0	1	x	x	x	x	x	x	x	x	x	x	x	0	0	0	0
0	0	1	x	x	x	x	x	x	x	x	x	x	0	0	0	1
0	0	0	1	x	x	x	x	x	x	x	x	x	0	0	1	0
0	0	0	0	1	x	x	x	x	x	x	x	x	0	0	1	1
0	0	0	0	0	1	x	x	x	x	x	x	x	0	0	1	0
0	0	0	0	0	0	1	x	x	x	x	x	x	0	0	1	1
0	0	0	0	0	0	0	1	x	x	x	x	x	0	0	1	0
0	0	0	0	0	0	0	0	1	x	x	x	x	0	0	1	1
0	0	0	0	0	0	0	0	0	1	x	x	x	0	1	0	0
0	0	0	0	0	0	0	0	0	0	1	x	x	0	1	0	1
0	0	0	0	0	0	0	0	0	0	0	1	x	0	1	1	0
0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1

Boolean functions:

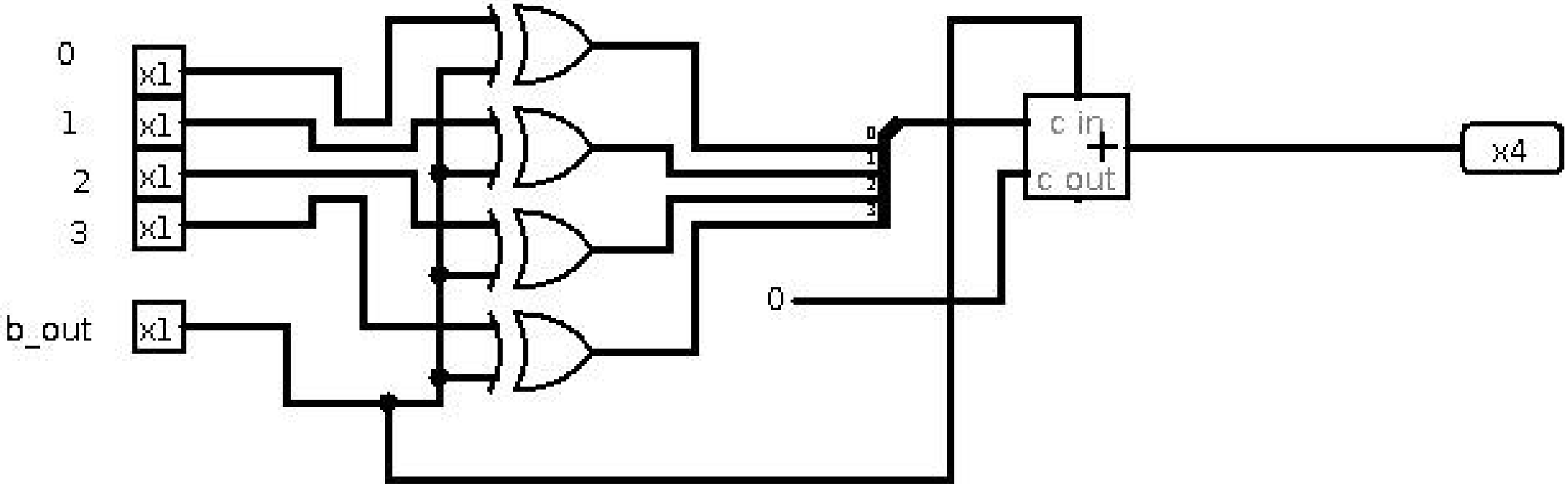
$$O_4 = I_{12}$$

$$O_3 = I_{12} + \overline{I_{12}}\,\overline{I_{11}}\,\overline{I_{10}}\,\overline{I_9}\,\overline{I_8}\,\overline{I_7}\,\overline{I_6}\,\overline{I_5}\,\overline{I_4}\,(I_3 + \overline{I_3}\,I_2 + \overline{I_3}\,\overline{I_2}\,I_1 + \overline{I_3}\,\overline{I_2}\,\overline{I_1}\,I_0)$$

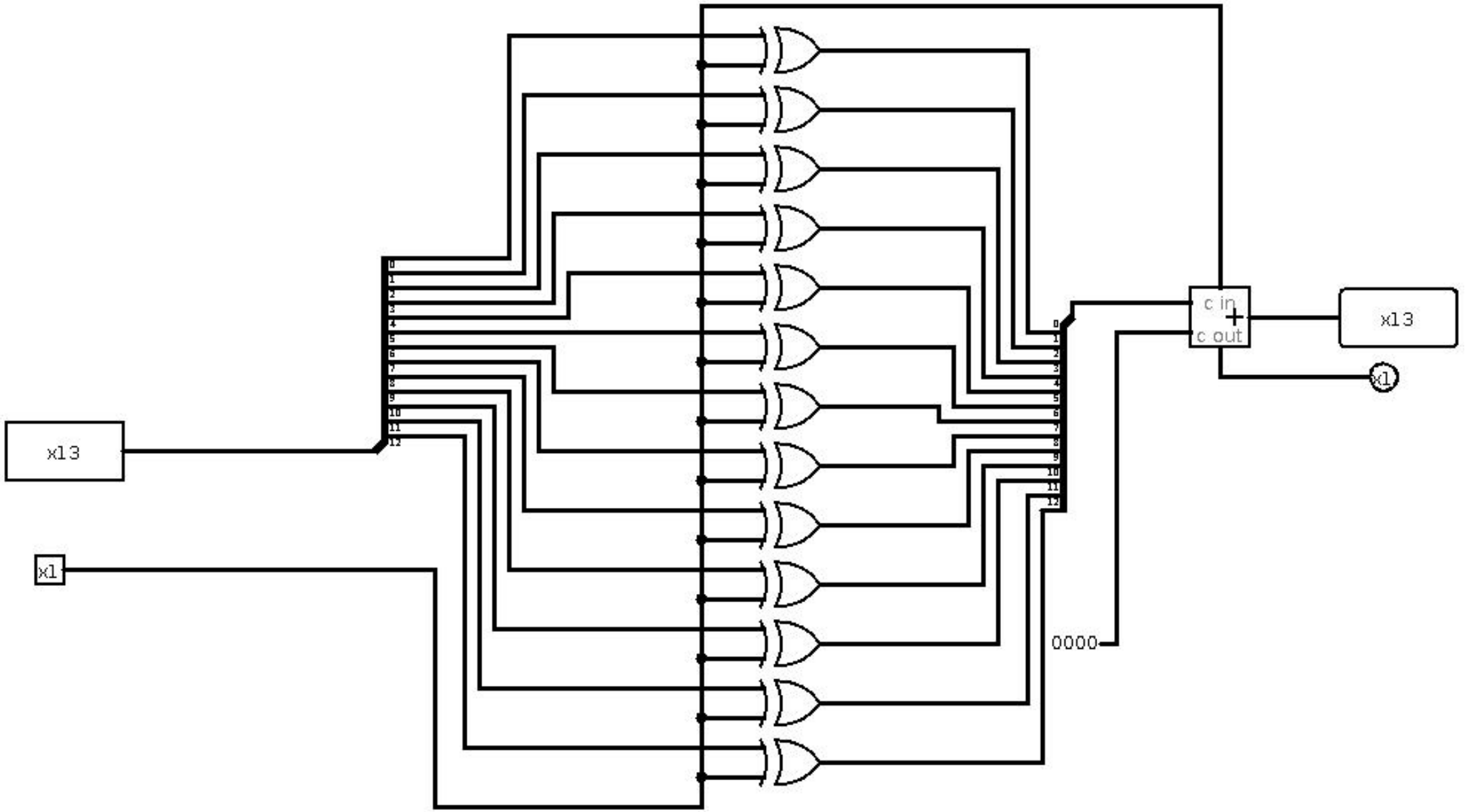
$$O_2 = I_{12} + \overline{I_{12}}\,\overline{I_{11}}\,\overline{I_{10}}\,\overline{I_9}\,\overline{I_8}\,(I_7 + \overline{I_7}\,I_6 + \overline{I_7}\,\overline{I_6}\,I_5 + \overline{I_7}\,\overline{I_6}\,\overline{I_5}\,I_4)$$

$$O_1 = I_{12} + \overline{I_{12}}\,\overline{I_{11}}\,\overline{I_{10}}\,(I_9 + \overline{I_9}\,I_8) + \overline{I_{12}}\,\overline{I_{11}}\,\overline{I_{10}}\,\overline{I_9}\,\overline{I_8}\,\overline{I_7}\,\overline{I_6}\,(I_5 + \overline{I_5}\,I_4) + \overline{I_{12}}\,\overline{I_{11}}\,\overline{I_{10}}\,\overline{I_9}\,\overline{I_8}\,\overline{I_7}\,\overline{I_6}\,\overline{I_5}\,\overline{I_4}\,\overline{I_3}\,\overline{I_2}\,(I_1 + \overline{I_1}\,I_0)$$

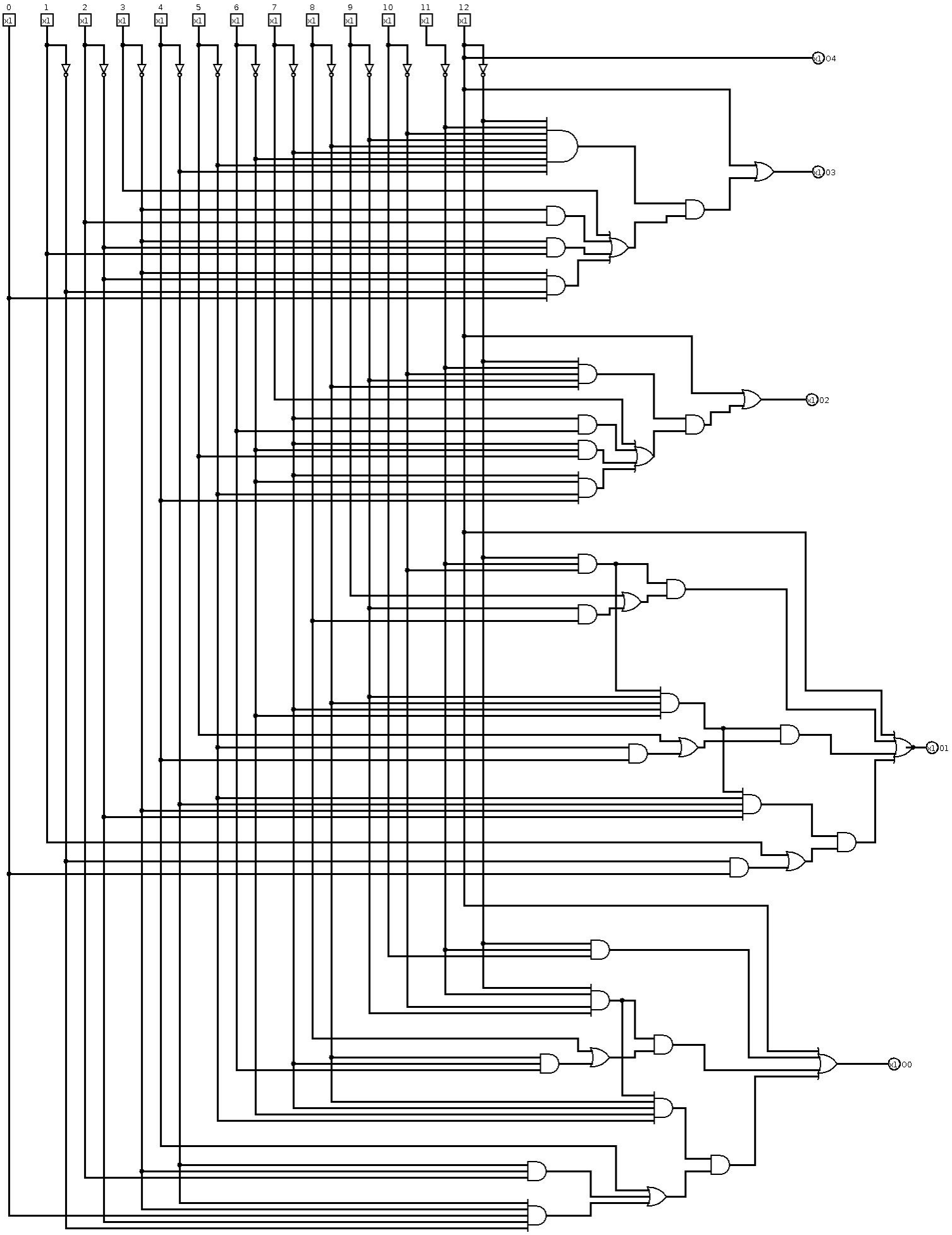
$$O_0 = I_{12} + \overline{I_{12}}\,\overline{I_{11}}\,I_{10} + \overline{I_{12}}\,\overline{I_{11}}\,\overline{I_{10}}\,\overline{I_9}\,(I_8 + \overline{I_8}\,\overline{I_7}\,I_6) + \overline{I_{12}}\,\overline{I_{11}}\,\overline{I_{10}}\,\overline{I_9}\,\overline{I_8}\,\overline{I_7}\,\overline{I_6}\,\overline{I_5}\,(I_4 + \overline{I_4}\,\overline{I_3}\,I_2 + \overline{I_4}\,\overline{I_3}\,\overline{I_2}\,I_1 + \overline{I_4}\,\overline{I_3}\,\overline{I_2}\,\overline{I_1}\,I_0)$$



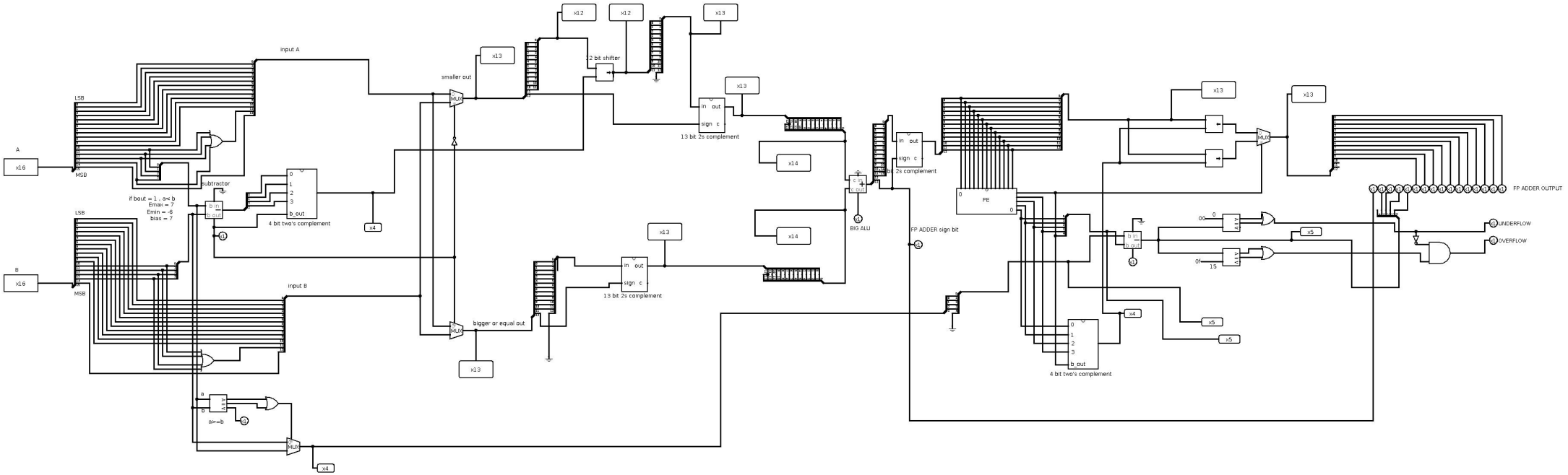
4bit 2's compliment



13bit 2's compliment



Priority Encoder



Floating Point Adder

IC's used with count as chart :

main circuit :

name	Count	IC no.
4-bit 2's compliment	2	
13-bit "	3	
Priority encoder	1	
Comparator	5	7485
MUX	10	74157
subtractor	2	
Adder	4	7483
Shifter	3	
Not	1	7404
OR	2	7432
AND	1	7408

4 - bit 2's compliment :

name	Count	IC NO.
XOR	1	7486
4bit full Adder	1	7483

13-bit 2's compliment:

name	count	IC NO.
XOR	~~13~~ 4	7486
13-bit Full Adder	4	7483

13-bit Priority Encoder:

name	count	IC NO.
Not	~~6~~ 2	7404
AND	~~27~~ 14	7408
OR	~~13~~ 5	7432

Simulator Used:

Logisim 2.7.1

Discussion :

The circuit we designed contains three sub-circuits we needed to build at first. Those are 4bit 2's compliment, 13 bit 2's compliment and a priority encoder. The 4 bit 2's compliment was used to convert the difference of exponents if it turns out negative. The 13 bit 2's compliment was used to convert the inputs if they were negative.

The priority encoder was designed to calculate how many bits output has to be shifted for normalisation along with its direction.

We performed the ~~fina~~ addition of fraction with a ~~bi~~ 14-bit adder passing 11-bit inputs with sign as msb along with the implicit 1. From the output of the adder, the msb is the sign bit of floating point addition, and first 13 bits are fraction bit which is normalised with priority encoder we built and ~~the~~ shifter.