Matlab arithmetic functions

- fix():Round toward zero

syntax : B = fix(A)

example : fix(-1.9) = -1

fix(5.6) = 5

- floor():Round toward negative infinity

syntax : B = floor(A)

example : floor(-1.9) = -2

floor(5.6) = 5

- round():Round to nearest integer

syntax : B = round(A)

example : round(-1.9) = -2

round(5.6) = 6

How to convert a number to a fixed point

f = number of decimal bits

n = floating point number

a = n * 2 ^ f

b = fix(a) or round (a)

binary fixed point = b / (2 ^ f)

Example

Given: n = 5.8515625 (in binary: 101 . 1101101)

Need to convert to fixedpoint with f = 3 decimal bits.

a = n * 2 ^ f = 5.8515625 * 2 ^ 3 = 46.8125008

b = fix(a) = fix (46.8125008) = 46

binary fixed point = b / (2 ^ f)

=46 / (2 ^ 3) = 5.75 (101 . 110)

Matlab Syntax example

```
f=3; %3 decimal bits
n=5.8515625;
fixed_point=(fix(n*2^3))/2^3;
```

Saturation

Considering the case that we have a 8-bits register and we are going to store a 9-bits number (1 0000 0000), if we store the first 8 bits the value of register would be zero which is wrong.

If our number exceeds the maximum value that can be represented, we should store the maximum value representable by the fixedpoint number rather that truncating data.

Matlab Example code

Number of required bits in our design: n=8 bits

Want to store a larger number e.g a = 512 ; % 1 0000 0000

```
if ( a > 2^(n-1)-1 )
    a_saturate = 2^(n-1)-1; % 0111 1111
else
    if( a < -2^(n-1) )
        a_saturate = -2^(n-1);
    end
end
```