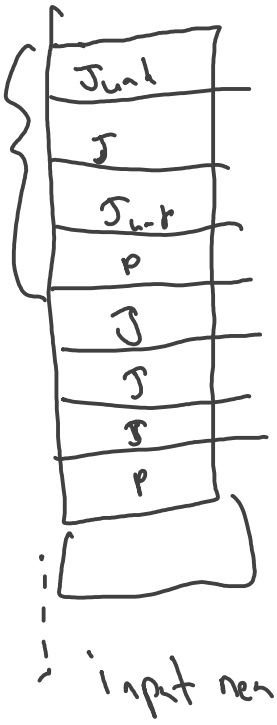
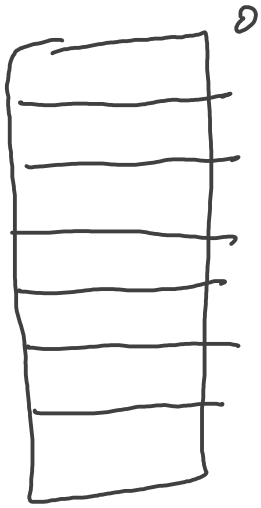
Notes on Digital Implementation

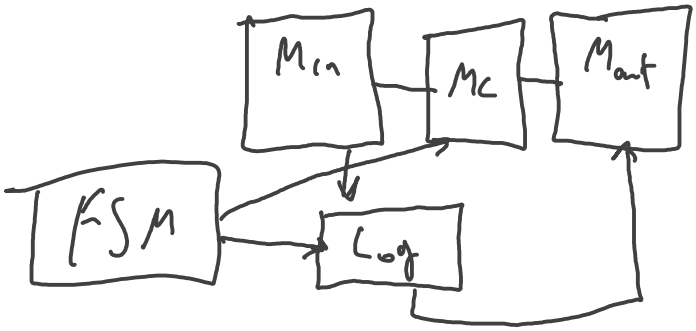
Thursday, October 17, 2024 10:17 AM

$Q = -10 \log_{10}(P)$ ← part of read info



1. Mem controller - keeps track mem addr, fet
2. * log module
3. Mean





Please:
1 take a look at the precision in terms of number of bits needed for log lookup table, identifying where the radix point is.

2. Look at the code to see if they're using a lookup table for the software implementation.
3. Take a look at what calculations are done after this.
4. Rename Conclusions section to Conclusions and Future Research

Phred scores are used to:

- Calculate per-base quality metrics, - what are these metrics? Are they calculated on each read? How are they calculated? Just the pscore or other items as well (from the read or other source)?
- Flag any low-quality bases for downstream filtering,
- Possibly compute average read quality or assess for high or low outliers.

Reads with overall low average scores or with individual bases below a set threshold can be flagged for potential filtering or further inspection.

After the initial quality checks, the processed read data can be passed on to modules for more detailed quality control, such as:

- Per-base sequence quality analysis,
- GC content and length checks,
- Analysis of overrepresented sequences or adapters if present.

Unfortunately, I overlooked the fact that the equation is not even used. The ASCII characters that are read are already calculated and we just need to convert the ASCII characters from the read.

FastQ Read Example:

@SEQ_ID GATCGGAAGAGCACACGTCTGAACTCCAGTCAC + IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII

1. The @ symbol is followed by a sequence Identifier which could be something like this:
@HWUSI-EAS100R:6:73:941:1973#0/1
2. Contains the nucleotide sequence and its bases consisting of A, T, G, C ex:
GATCGGAAGAGCACACGTCTGAACTCCAGTCAC
3. The + symbol is just a separator in the read
4. IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII represents ascii code for the quality score

@HWUSI-EAS100R:6:73:941:1973#0/1 NAGCTCGTTCGATGATCCTG + #%%&''()*+,-./01234