~~a(b|ab)(a?ab)*~~

1) NFSA to Regular Expression

(a) Lets ~~write~~ write the possible ~~solution~~ ways

(1) ab

(2) aab

(3) abab

(4) abaab

(5) aababaab

There can be many repeating combinations

RE =

$$a(b|ab)(a?ab)^*$$

(b) possible strings (combinations)

(1) ab

(2) aba

(3) ababababab ---

(4) abaaab

(5) abaaba

There may be more possibilities

RE $\boxed{a(ba|baa)^*(ba|b)}$

2) Sentence

The standard turbo engine is hard to work

1) Prob

$= P(\text{standard} \mid \text{The}) * P(\text{turbo} \mid \text{standard}) * P(\text{engine} \mid \text{turbo})$

$* P(\text{is} \mid \text{engine}) * P(\text{hard} \mid \text{is}) * P(\text{to} \mid \text{hard}) *$

$P(\text{work} \mid \text{to})$

$= 0.0008161 \times 0.2 \times 0 \times 0 \times 0 \times 0.75 \times 0.004513$

$= 0 \Rightarrow$ No smoothing

2) Add-one (Bigrams)

$= 0.000354 \times 0.000394 \times 0.0001315 \times 0.0001312$

$* 0.000124 * 0.000525 * 0.0008740$

$= 1.3778 \times 10^{-25} \Rightarrow$ Add one smoothing

III Good Turing (Bigram)

$= (4.5209 \times 10^{-5}) * (1.2824 \times 10^{-5}) * (0.248357)$

$* (0.248357) * (0.248357) * (4.520505 \times 10^{-5})$

$* 0.0001794$

$= 4.33453 \times 10^{-20} \rightarrow$ Good Turing

For Unigrams (for the given sentence)

$= P(\text{The}) * P(\text{standard}) * P(\text{Turbo}) * P(\text{engine})$

$* P(\text{is}) * P(\text{hard}) * P(\text{to}) * P(\text{work})$

$= 0.053492 * 0.00145 * 2.109641 * 0.0024 *$

$0.00650 * 5.819 \times 10^{-5} * 0.02256 * [illegible]$

$* 0.0005814$

$= 2.782076 \times 10^{-25} \Rightarrow$ for Unigrams