Viterbi Algo :-

- most common decoding algo.
- Decoding task - determine the sequence of variables of the underlying source of some sequence of observation.
- It is a dynamic programming soln.

Viterbi algo.

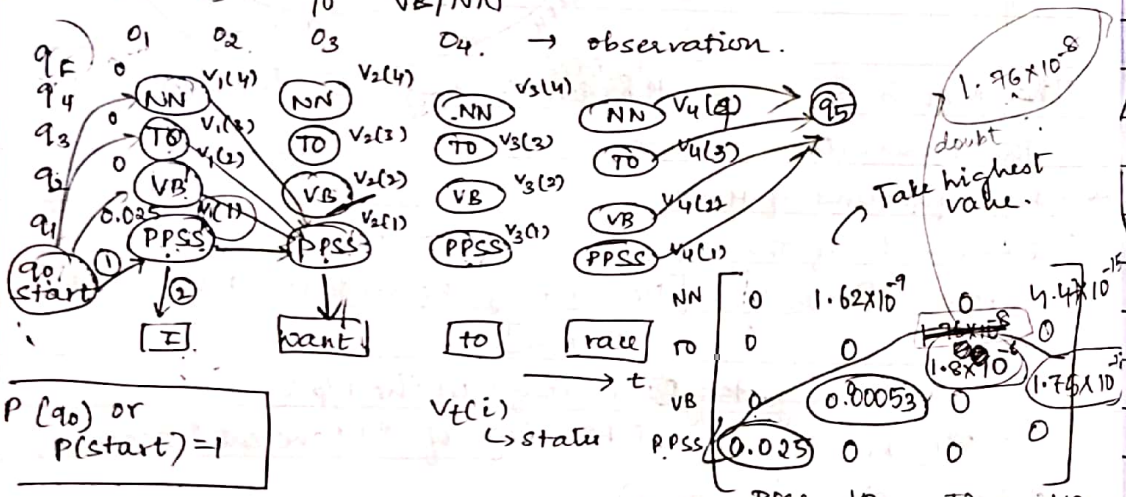
I/p : a single HMM and a set of observed words $O = (o_1 o_2 o_3 \dots o_T)$

Returns : the most probable state / tag sequence $Q = (q_1 q_2 \dots q_T)$ together with its probability.

Eg: "I want to race" → I/p

PP VB TO VB/NN

o_1 o_2 o_3 o_4 → observation.



P(q0) or P(start) = 1

① ⇒ P(PPSS | <s>) = 0.067.

② ⇒ P(I | PPSS) = 0.037

∴ $V_1(1) = P(\text{PSS} \mid \langle s \rangle) \cdot P(\text{I} \mid \text{PPSS}) = 0.067 \times 0.037$
$= 0.025$

P(start)

$V_1(2) = P(\text{VB} \mid \langle s \rangle) \cdot P(\text{I} \mid \text{VB}) = 0.$

$V_1(3) = P(\text{TO} \mid \langle s \rangle) \cdot P(\text{I} \mid \text{TO}) = 0$

$V_1(4) = P(\text{NN} \mid \langle s \rangle) \cdot P(\text{I} \mid \text{NN}) = 0.$

$P(\text{NN} \mid \langle s \rangle) \cdot P(\text{I} \mid \text{NN})$

0.067
0.37
0.025

Tag Transition prob.

	VB	TO	NN	PPSS
<s>	0.019	0.0043	0.041	0.067
VB	0.0038	0.035	0.047	0.0070
TO	0.83	0	0.00047	0
NN	0.0040	0.016	0.087	0.0045
PPSS	0.23	0.00079	0.0012	0.00014

Observation likelihood

	I	want	to	race
VB	0	0.0093	0	0.00012
TO	0	0	0.99	0
NN	0	0.00054	0	0.00057
PPSS	0.37	0	0	0.

Compute $V_2(1) = \max\{\text{tag} \cdot P(\text{want} \mid \text{PPSS})\}$.

$= \max\{P(\text{PPSS}) * P(\text{PPSS} \mid \text{PPSS}), P(\text{VB}) * P(\text{PPSS} \mid \text{VB}), V_3(1) \in P(\text{TO}) * P(\text{PPSS} \mid \text{TO}), P(\text{NN}) * P(\text{PPSS} \mid \text{NN})\} \cdot P(\text{want} \mid \text{PPSS})$

$V_2(1)$ → 0, $V_4(1)$, 0

$= 0.025 * 0.00014 * 0 = 0$

$V_2(2) = \max\{V_1(2) \cdot P(\text{VB} \mid \text{VB}), V_1(1) \cdot P(\text{VB} \mid \text{PPSS}), V_1(3) \cdot P(\text{VB} \mid \text{TO}), V_1(4) \cdot P(\text{VB} \mid \text{NN}) * P(\text{want} \mid \text{VB})\}$

0, 0

$0.23 * 0.025 * 0.0093$

$= 0.000053$

$V_2(3) = \max\{V_1(1) \cdot P(\text{TO} \mid \text{PPSS}), V_2(2) \cdot P(\text{TO} \mid \text{VB}), V_2(3) \cdot P(\text{TO} \mid \text{TO}), V_1(4) \cdot P(\text{TO} \mid \text{NN})\} \cdot P(\text{want} \mid \text{TO})$

$= \{0.025 * 0.00079, 0, 0, 0\} * 0 = 0$

HW :-

Eg 2: learning changes thoroughly'

	NN	VB	Adv	qF
q_0	2×10^{-1}	3×10^{-1}	0	0
NN	1×10^{-1}	3×10^{-1}	1×10^{-1}	0
VB	4×10^{-1}	1×10^{-1}	4×10^{-1}	0
Adv	0	0	0	1×10^{-1}

	learning	changes	thoroughly
NN	1×10^{-3}	3×10^{-3}	0
VB	3×10^{-3}	4×10^{-3}	0
Adv	0	0	2×10^{-3}