State transition table below:

To **From**	S	X	NC	E
S	0	0,5	0,5	0
X	0	0,8	0,1	0,1
NC	0	0,2	0,6	0,2
E	0	0	0	1

Probability to generate one of the four nucleotide subunits A, C, G, T in the four states is given in the table below:

Nucleotide generation **State**	A	C	G	T
S	0	0	0	0
X	0,4	0,1	0,2	0,3
NC	0,4	0,1	0,1	0,4
E	0	0	0	0

a. A DNA chain starting with two nucleotides in the following order ‘G’, ‘T’ is generated. What is the most likely sequence of states to have generated this start of a DNA chain? Use the Viterbi algorithm.
Hint: you may ignore the end state (E) in the calculations.

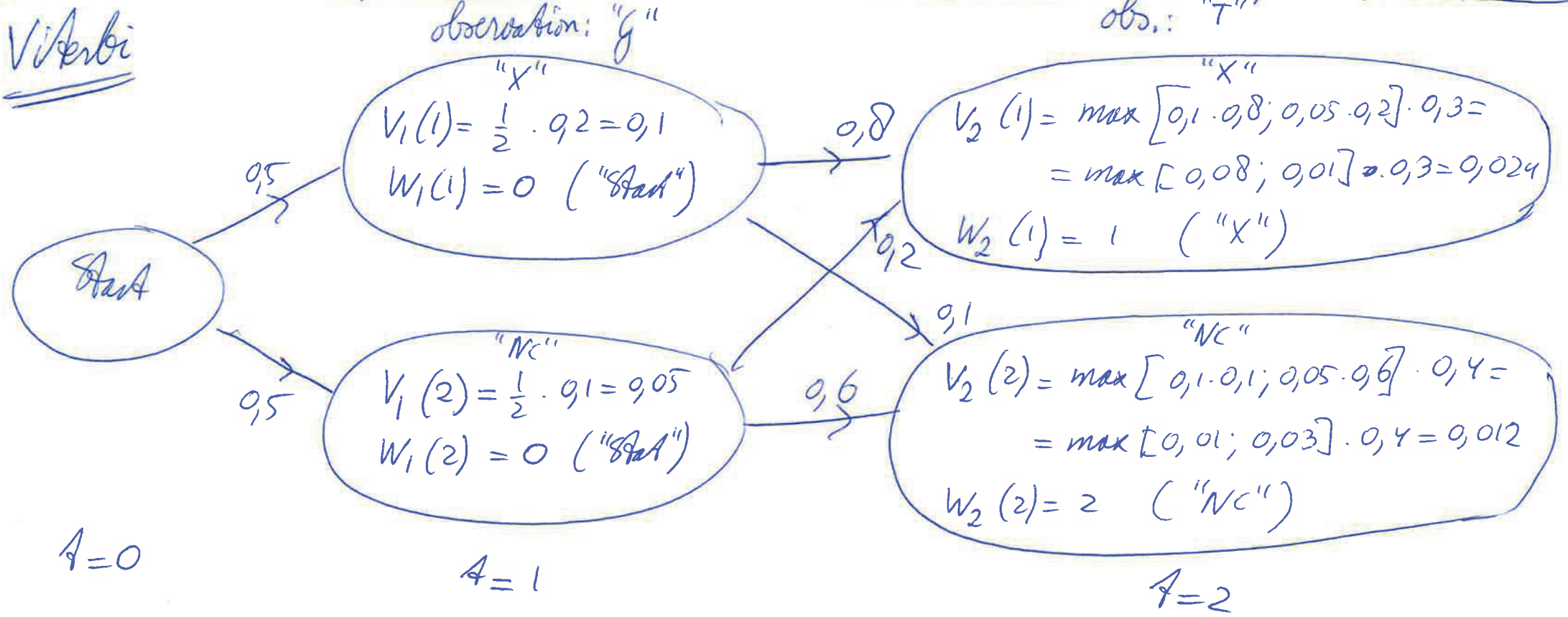
(1d) Notation: States i: $i=0$: "Start" $i=1$: "X" $i=2$: "NC"

$V_t(i)$: probability of <u>best path</u> ending in state i at time t.

↓ sequence of states that generates the required observations ("G", "T")

$W_t(i)$: state at time $t-1$ of best path corresponding to $V_t(i)$

Viterbi



ANSWER: $V_2(1) > V_2(2)$: best path ends at time $t=2$ in "X": $s_2^* = 1$

sequence of states generating ("G", "T"): (s_0^*, s_1^*, s_2^*)

$$s_1^* = w_2(s_2^*) = w_2(1) = 1 \quad (\text{"X"})$$

$$s_0^* = w_1(s_1^*) = w_1(1) = 0 \quad (\text{"Start"})$$

Thus: most likely sequence of states to generate "G", "T":

"Start" → "X" → "X"