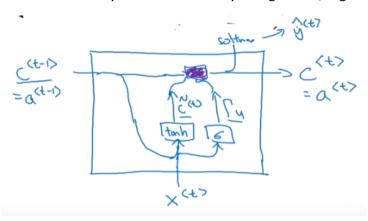
	1	2	3	4	5	6	7	8	9	10
c ~ (memory candidate)	•••		1			•••	•••	•••	0	•••
gamma (if to update)	0	0	1	0	0	0	0	0	1	0
c (memory)	0	0	1	1	1	1	1	1	0	0
Х	 	The	cat	which	already	ate	•••	•••	was	full
y (output with softmax)	The	cat	which	already	ate	•••	•••	was	full	<end of="" sentence=""></end>
a (output w/o softmax)	•••									•••

- > Sequence to Sequence, vanishing gradients
- > Can have many different memory and gamma, e.g. ate -> full



Step-by-step example

Step 1	1	2	3	4	5	6	7	8	9	10
c ~ (memory candidate)										
gamma (if to update)	0	0								
c (memory)	0	0								
X	 	The	cat							
y (output with softmax)	The	cat								
a (output w/o softmax)										

Step 2	1	2	3	4	5	6	7	8	9	10
c ~ (memory candidate)										
gamma (if to update)	0	0	1							
c (memory)	0	0								
X	 degin of sentence>	The	cat							
y (output with softmax)	The	cat								
a (output w/o softmax)										
Step 3	1	2	3	4	5	6	7	8	9	10
c ~ (memory candidate)										
gamma (if to update)	0	0	1							
c (memory)	0	0	1							
х	 	The	cat							
y (output with softmax)	The	cat	which							
a (output w/o softmax)										