

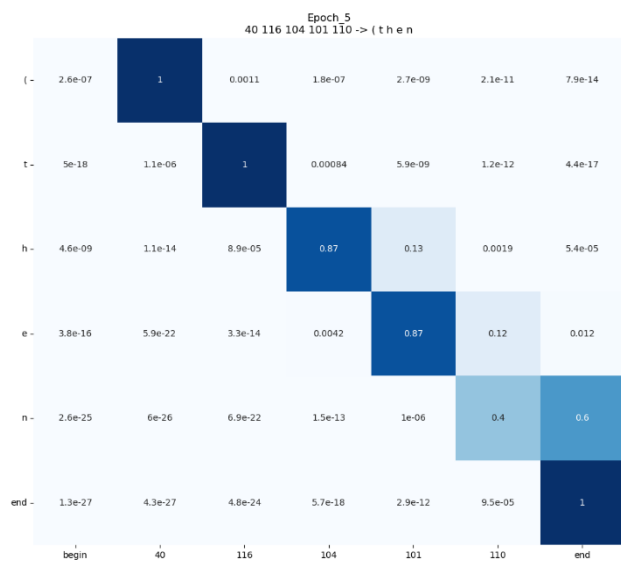
## Ex2 – Part 3 – Report

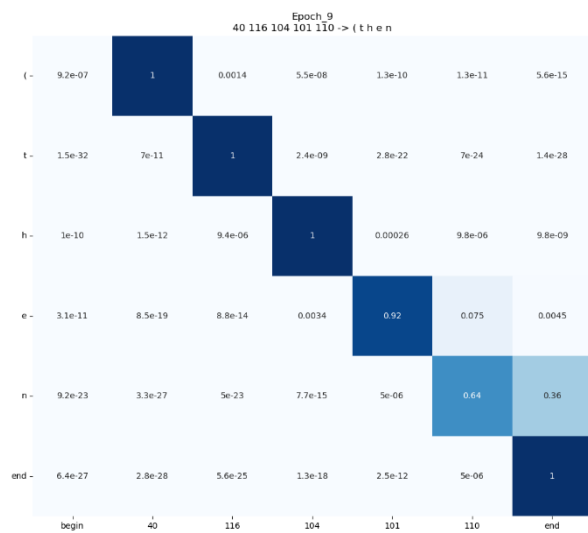
Epoch_1 40 116 104 101 110 -> ( t h e n							
f -	9.4e-24	3.5e-21	1.8e-35	7.1e-37	1.5e-23	3.5e-24	1
t -	3e-22	3.6e-20	1.1e-29	1.6e-30	1.7e-20	3.5e-20	1
h -	1.9e-20	9.2e-19	1.5e-26	3.6e-27	1.5e-18	4.9e-18	1
e -	1.3e-19	5.2e-18	2.4e-25	5.8e-26	9.7e-18	3.5e-17	1
n -	1.7e-18	5e-17	1e-23	2.9e-24	1.3e-16	4.9e-16	1
end -	7.7e-19	2.2e-17	3.4e-24	9.7e-25	4.7e-17	1.8e-16	1
	begin	40	116	104	101	110	end

Epoch_2 40 116 104 101 110 -> ( t h e n							
f -	1.1e-21	3.2e-21	2e-30	7.8e-34	1.7e-13	1.6e-21	1
t -	2.4e-20	4e-19	1.5e-21	1.4e-23	2e-12	1.1e-14	1
h -	5.3e-13	1.9e-12	6.9e-13	5.4e-14	9.8e-09	3.3e-09	1
e -	2.3e-11	6.4e-11	2.9e-11	3.3e-12	8.5e-08	3.8e-08	1
n -	4.4e-10	1e-09	5.2e-10	7.6e-11	4.6e-07	2.5e-07	1
end -	8.8e-10	2e-09	1e-09	1.6e-10	7.3e-07	3.9e-07	1
	begin	40	116	104	101	110	end

Epoch_3 40 116 104 101 110 -> ( t h e n							
f -	2.1e-07	1	6.1e-05	1.4e-09	6.1e-11	6.4e-14	5.4e-20
t -	3.1e-28	1	5.7e-07	4.1e-21	2.9e-11	2.3e-18	3.2e-33
h -	4.9e-22	2e-13	3.5e-07	8.8e-10	0.26	0.68	0.056
e -	7.4e-21	4e-19	9.8e-18	2.7e-18	5.6e-13	2e-08	1
n -	8.7e-15	1.5e-13	6.8e-13	3.3e-13	1.6e-10	9.1e-08	1
end -	1.3e-15	1.7e-14	1.1e-13	5.1e-14	6.5e-11	8.3e-08	1
	begin	40	116	104	101	110	end

Epoch_4 40 116 104 101 110 -> ( t h e n							
f -	2.8e-06	1	2.2e-05	7.9e-08	1.7e-09	3.2e-12	4.9e-15
t -	3e-19	1.4e-08	1	0.0026	0.00023	2.6e-11	2.7e-18
h -	3.3e-12	3e-15	0.0033	0.5	0.5	0.0022	2.5e-05
e -	1.3e-20	3.3e-22	4e-12	1.8e-05	0.46	0.52	0.014
n -	1.3e-25	8e-24	7.8e-20	2.6e-16	4.1e-08	0.17	0.83
end -	1.2e-24	3.4e-22	6.6e-20	2.3e-18	1.7e-13	1.5e-06	1
	begin	40	116	104	101	110	end





We can see that at the beginning of the training the attention is mostly given to the end token ( $\langle /s \rangle$ ) – this occurs because at the first steps our context vector doesn't know how to align the words in the sentence.

As the training makes progress, we can see that the attention converges towards the diagonal of the heatmap – each token gets its appropriate attention to the translation.