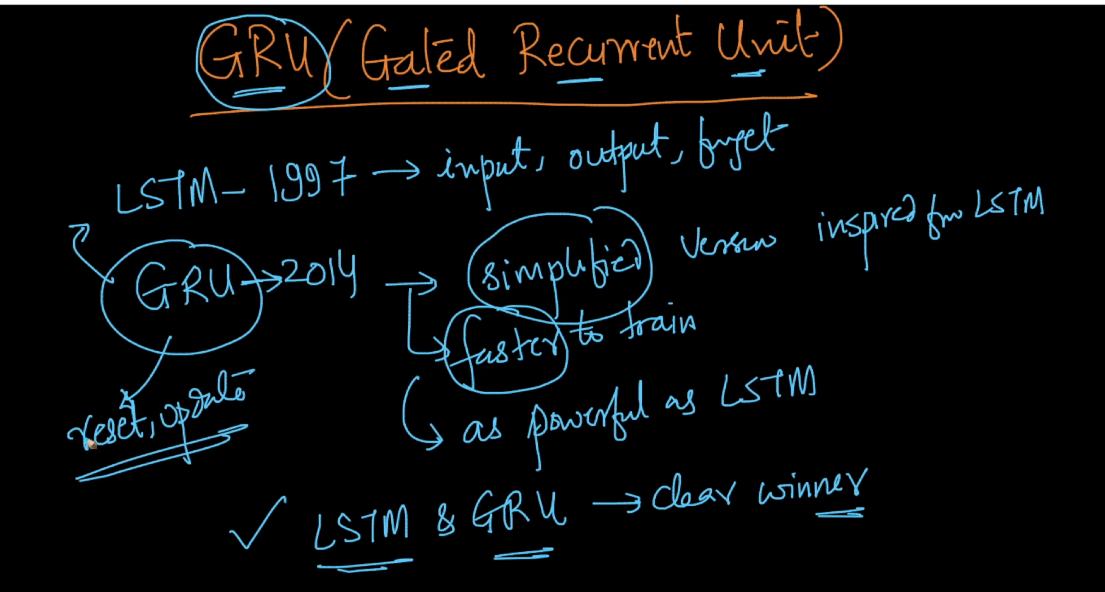
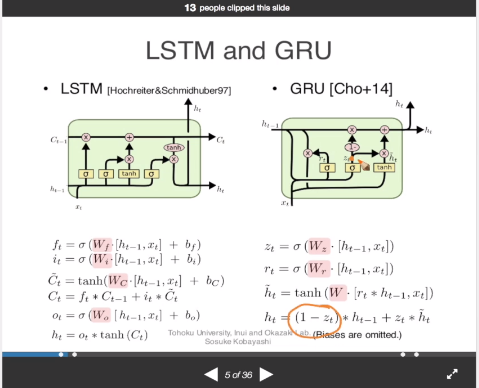
**GRUs**

GRU is a simplified version of LSTM, it have less derivative and less structure only have reset and update structure unlike LSTM that have input, output, forget structure. Therefore it is faster to train still it is as powerful as LSTM.



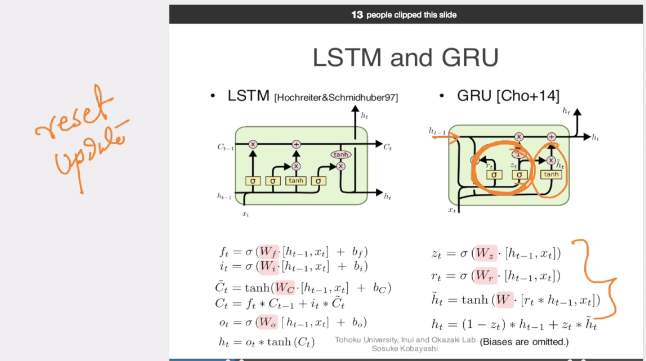
Below shows structure of LSTM and GRU both here we can see that GRU is very simple and have less derivative

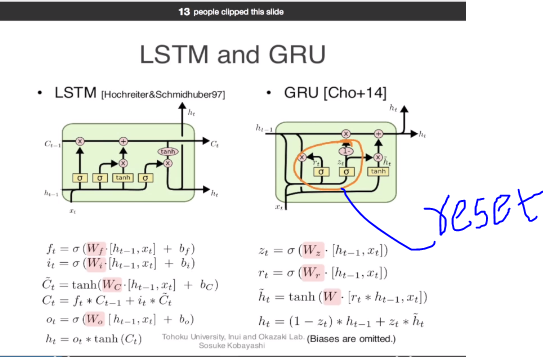


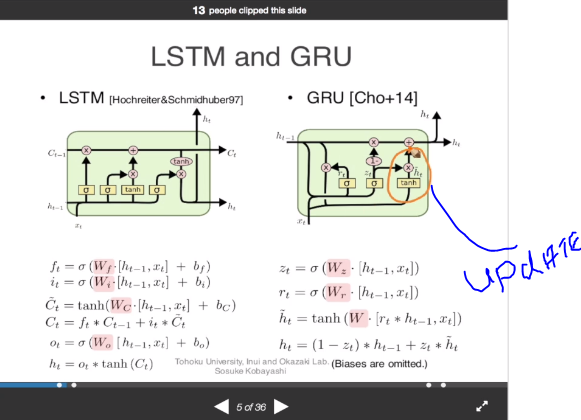
Below image shows reset and update structure.

So what happen is as seen in first equation if zt is 1 then by eq4 of ht we include current input and neglect previous output.

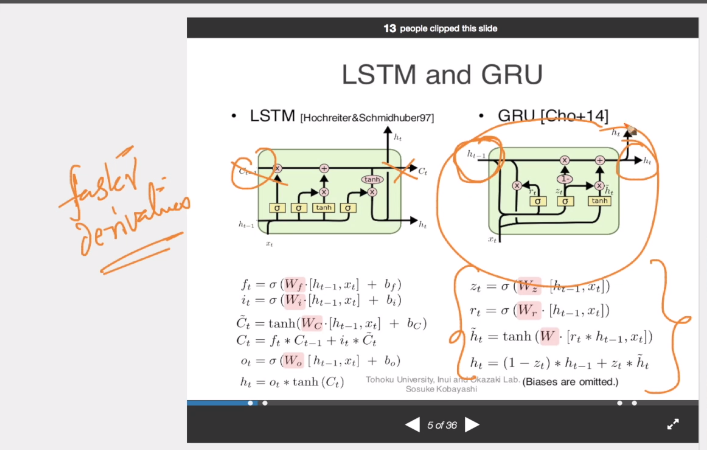
And if zt is 0 then by eq4 we include previous output.



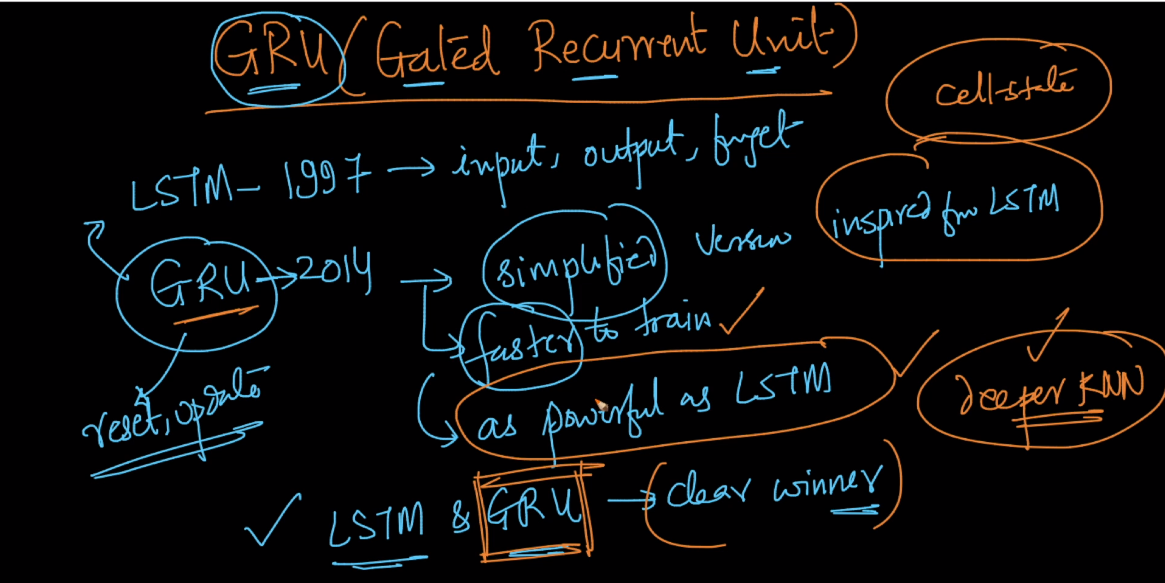




Derivatives of GRU are faster.



Till now we see only single layer RNN but we can use GRU for deeper RNN as well.



Link :

Very good video by Andrew ng : <https://www.youtube.com/watch?v=xSCy3q2ts44>

<http://colah.github.io/posts/2015-08-Understanding-LSTMs/>

<https://www.slideshare.net/hytae/recent-progress-in-rnn-and-nlp-63762080>