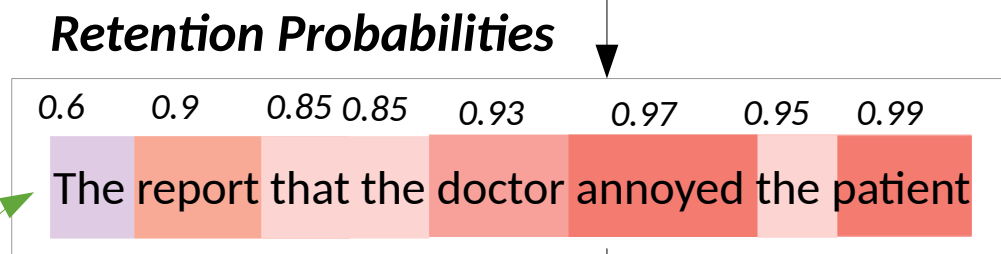


True Context c^*

The report that the doctor annoyed the patient...



Lossy Representation $c' \sim P(c'|c^*)$

?? report ?? ?? doctor annoyed the patient...

Posterior $P(c|c') \propto P(c'|c) P(c)$

The report that the doctor annoyed the patient...



The report by the doctor annoyed the patient...



A report about the doctor annoyed the patient...



...

Bayesian Inference

Next-Word Predictions $P(w|c)$

was
came
yesterday
End of Sentence
...



was
came
yesterday
End of Sentence
...



was
came
yesterday
End of Sentence
...



**Next-Word Prediction
 $P(w|c') = \sum_c P(w|c) P(c|c')$**

was
came
yesterday
End of Sentence
...



**Surprisal
 $-\log P(w|c')$**

Training Signal

optimize retention probabilities