Experimental and Computational Methods in Linguistic Research

Spring 2025

Instructor: Sanghee Kim

Week 7

Agenda

- Reading time and surprisal
- Language model output

- Research proposal
  - Research question
  - Hypothesis
  - Experimental design
  - Methods
  - Analysis
  - Predictions
  - Report

Final presentation

- Sign up for your final presentation (May 20 or May 22)

Understanding reading times

- Why do we see such reading time differences?

The debate

- Memory?
- Expectation?

https://rlevy.github.io/9.19-syllabus/assets/slides/pcfgs-garden-pathing-and-surprisal-with-builds.pdf

Surprisal & Psycholinguistics

- In addition to measuring the average information for a language, we can of course measure the **information conveyed by any given linguistic unit** (e.g. phoneme, word, utterance) in context. This is often called *surprisal*:

$$Surprisal(x) = \log_2 \frac{1}{P(x \mid context)}$$

- **Surprisal will be high**, when *x* has a low conditional probability, and **low**, when *x* has a high probability.

- Claim: **Cognitive effort** required to process a word is **proportional** to its **surprisal** (Hale, 2001).

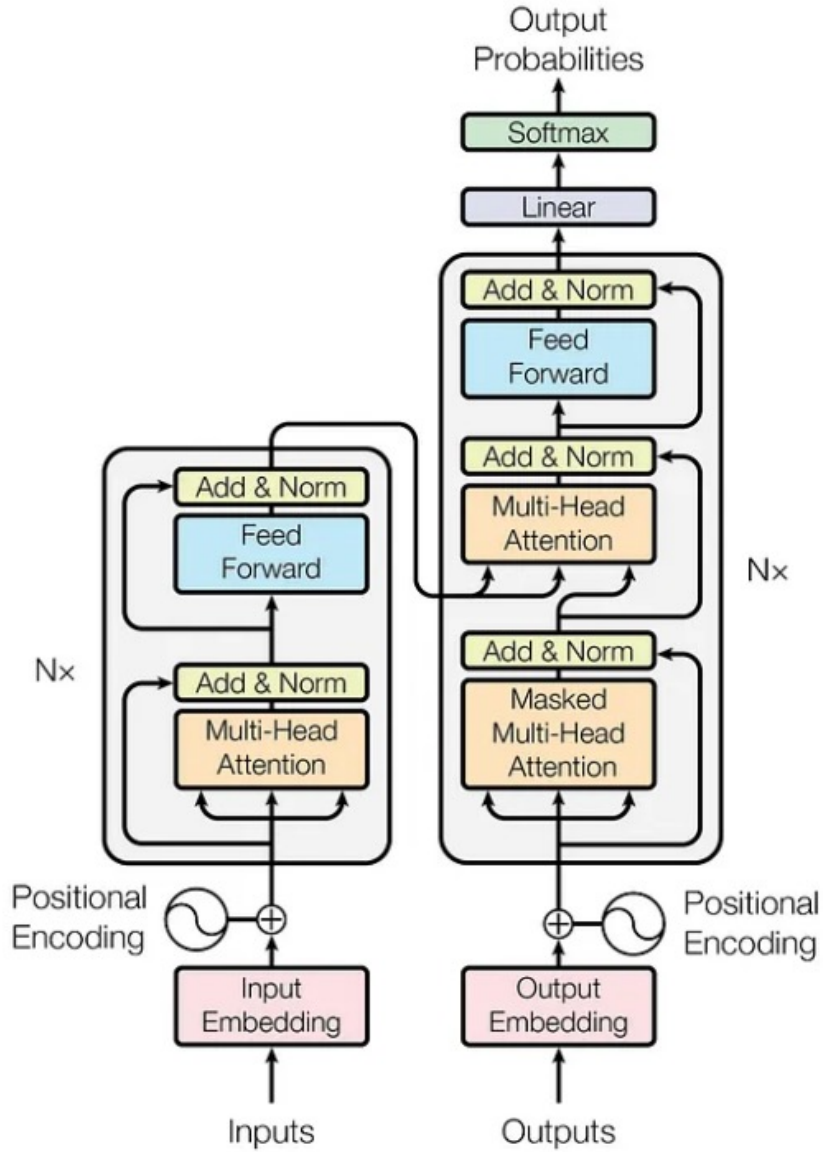
Surprisal and log probability

$$Surprisal(x) = \log_2 \frac{1}{P(x \mid context)}$$

$$\textbf{\textit{Surprisal}(x)} = \textbf{-log}_2\textbf{P(x)}$$

BERT

Encoder

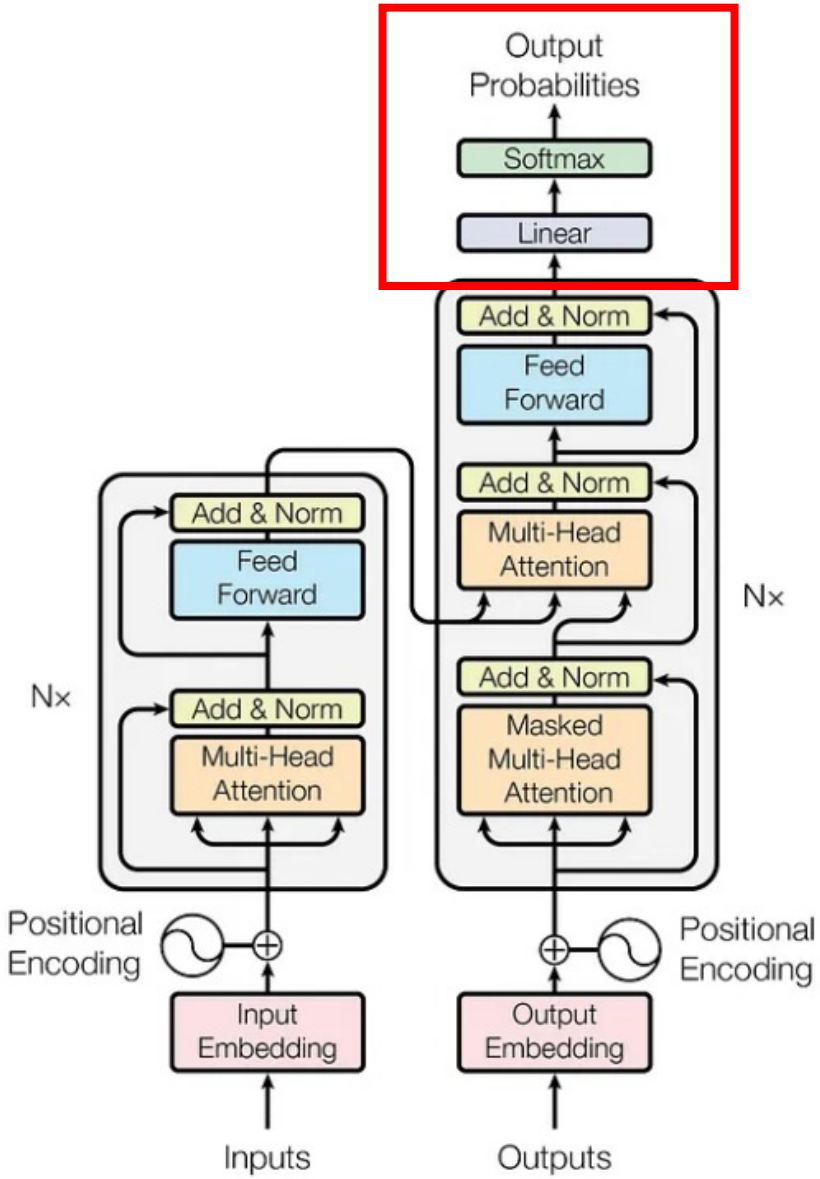


Transformer Architecture

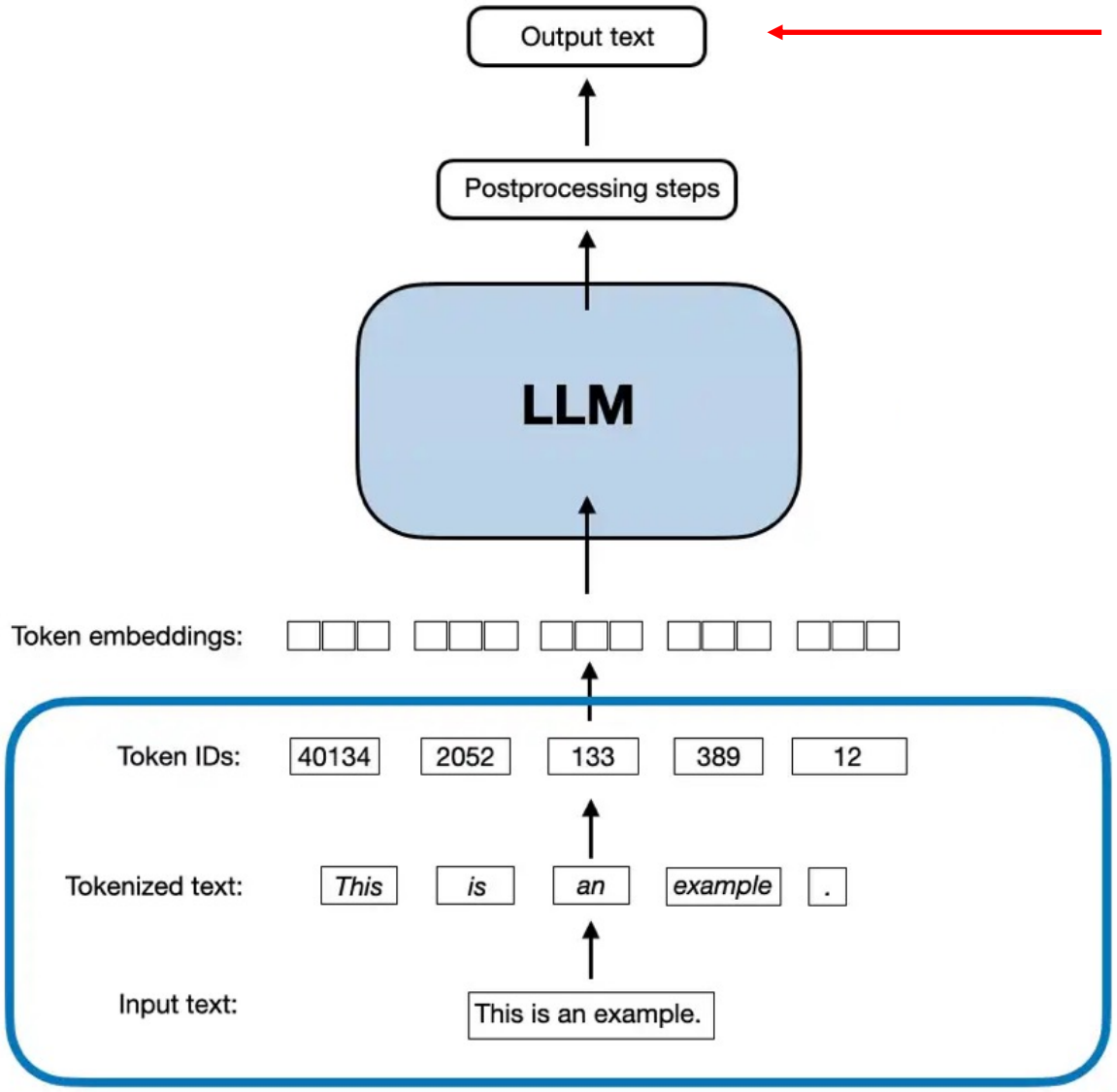
GPT

Decoder

- Probability distribution over words



Transformer Architecture



We will access the model output to obtain the probability (and surprisal) of the word (given prior context).

https://sebastianraschka.com/blog/2025/bpe-from-scratch.html

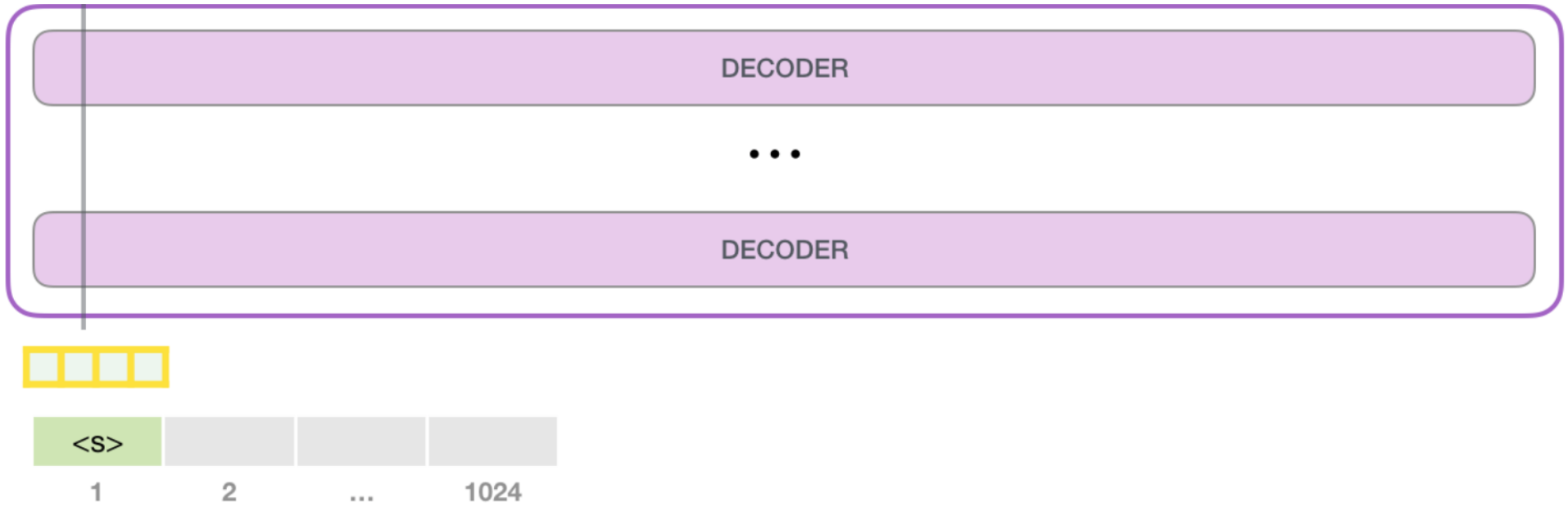
Input

https://jalammar.github.io/illustrated-gpt2/

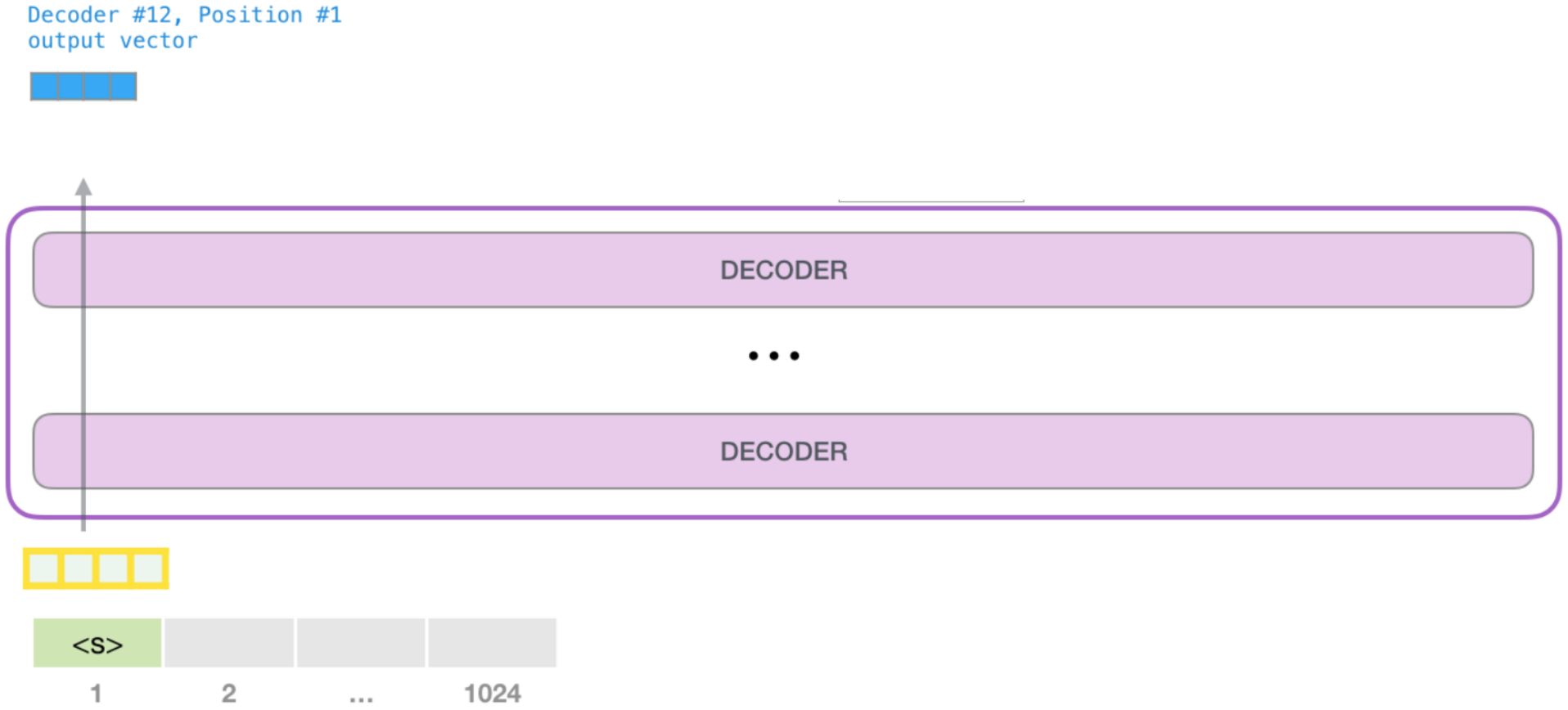
Input tokenization

https://jalammar.github.io/illustrated-gpt2/

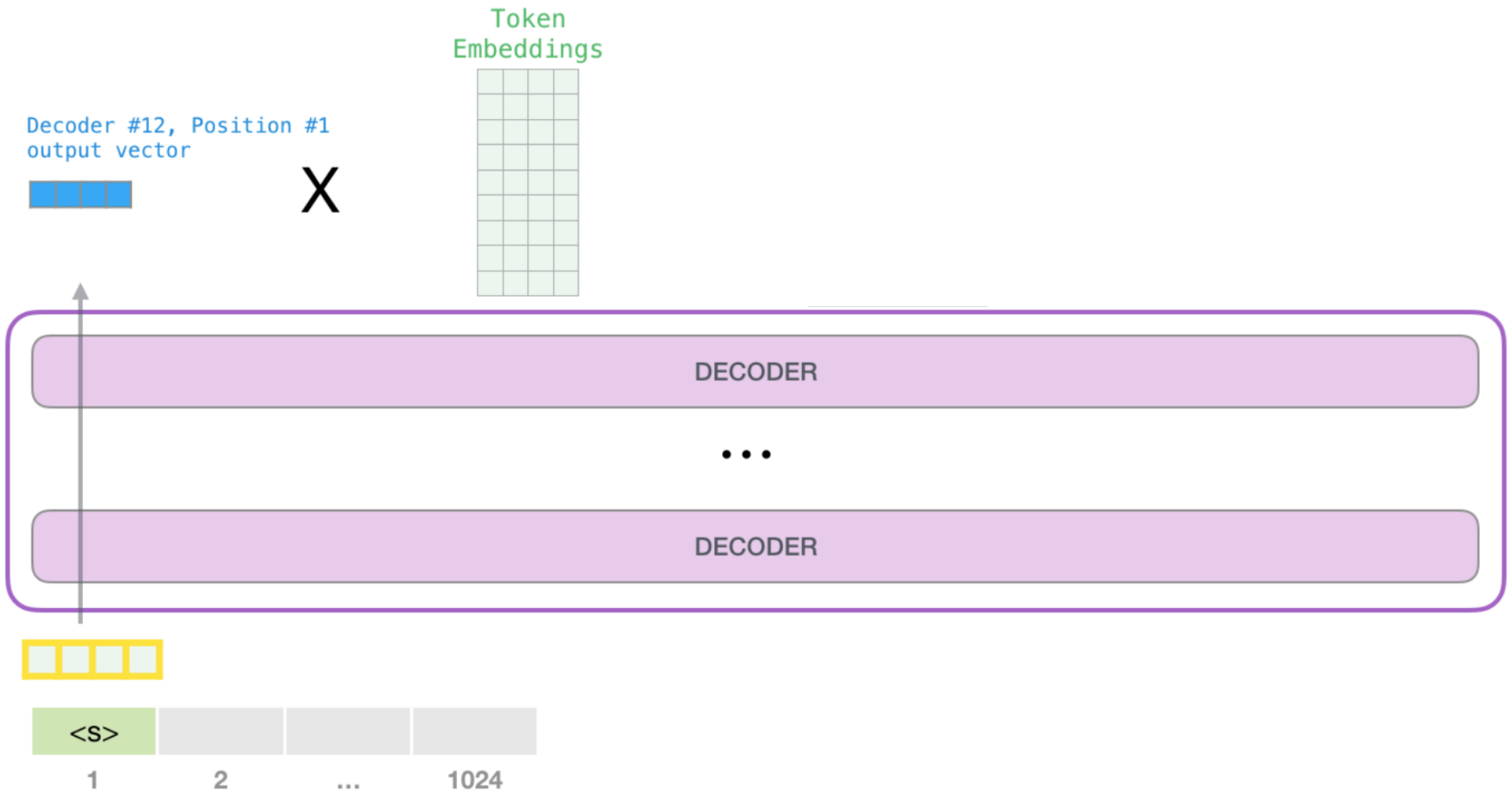
https://jalammar.github.io/illustrated-gpt2/



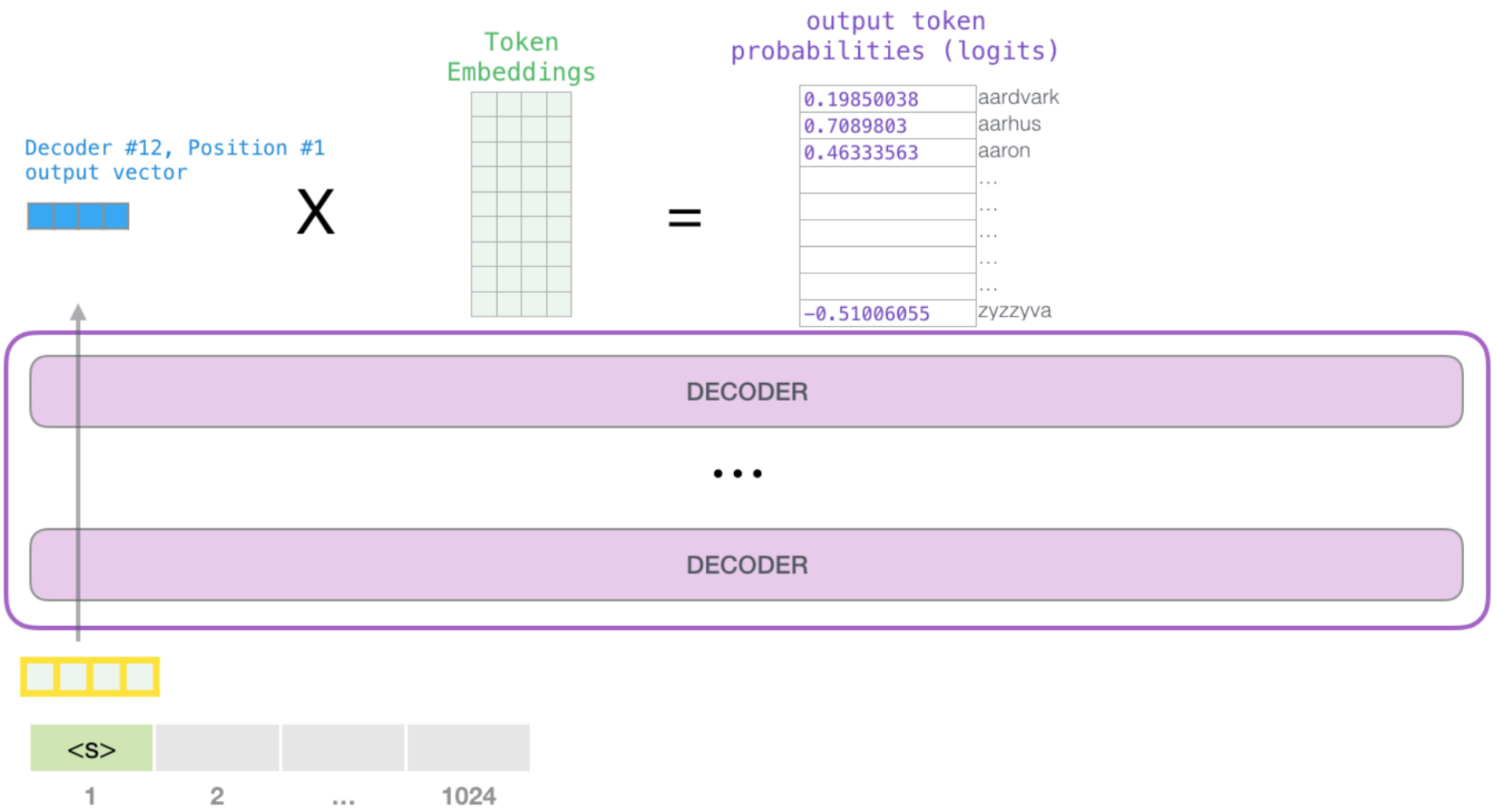
https://jalammar.github.io/illustrated-gpt2/



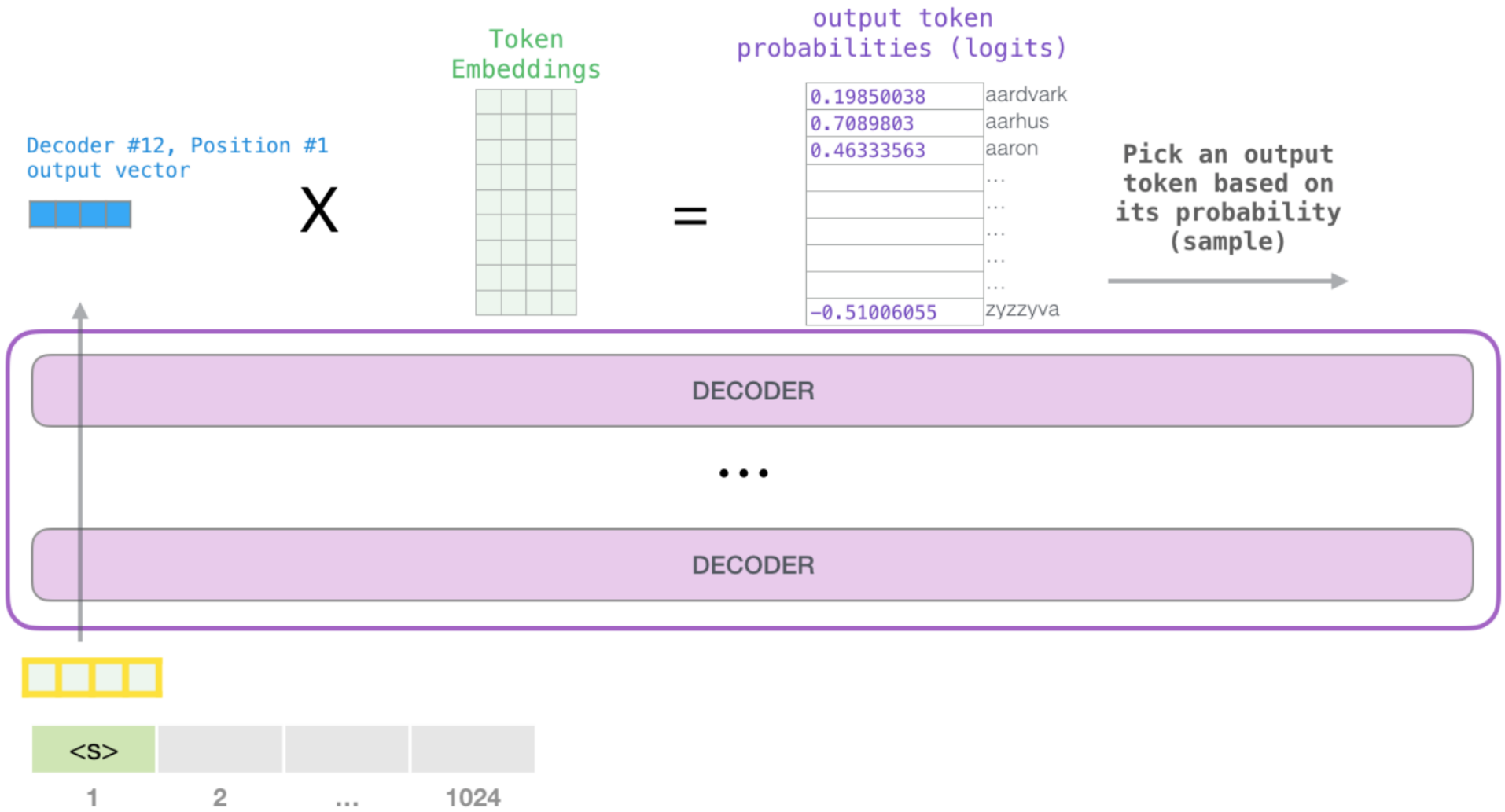
https://jalammar.github.io/illustrated-gpt2/



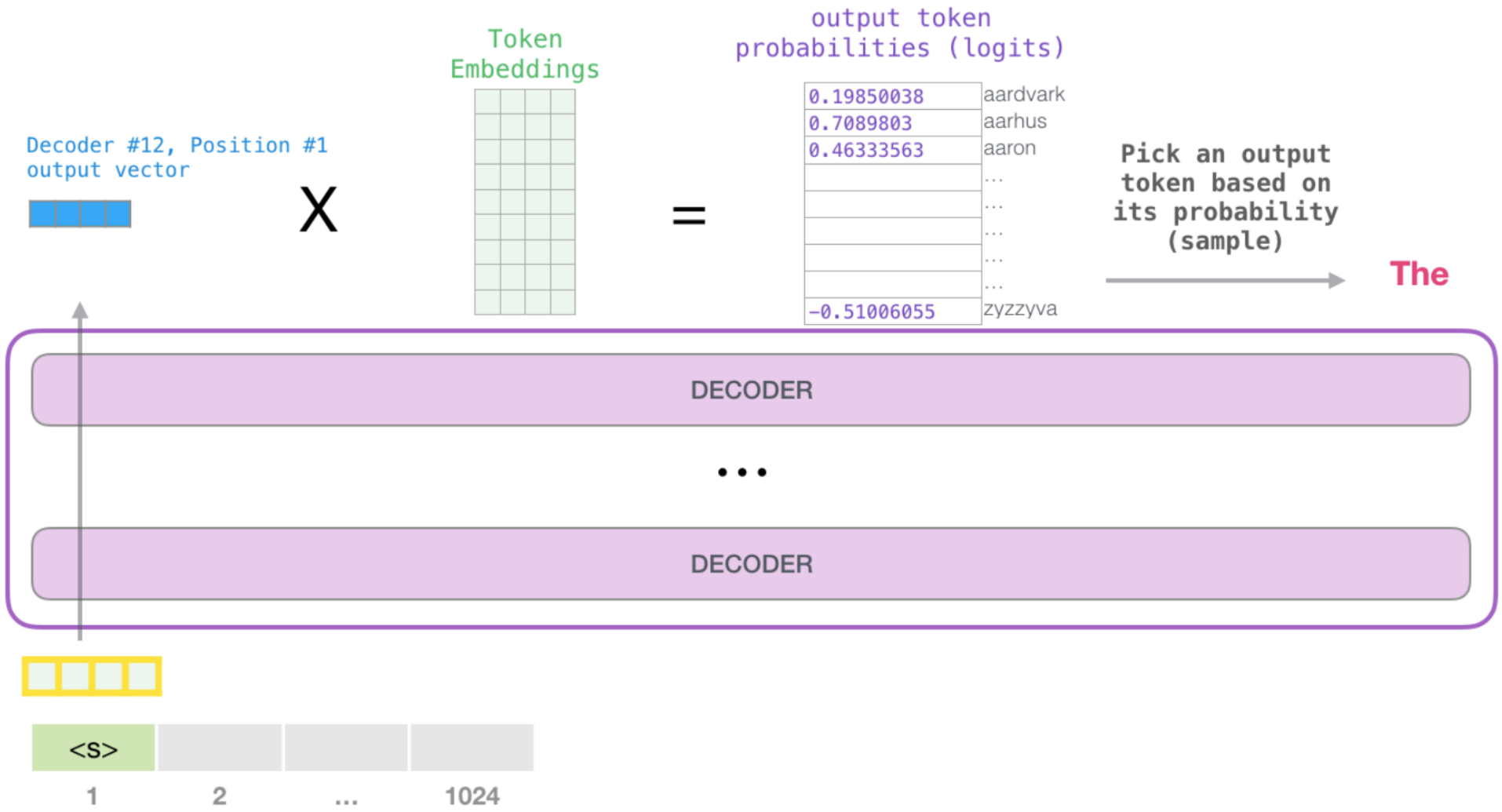
https://jalammar.github.io/illustrated-gpt2/



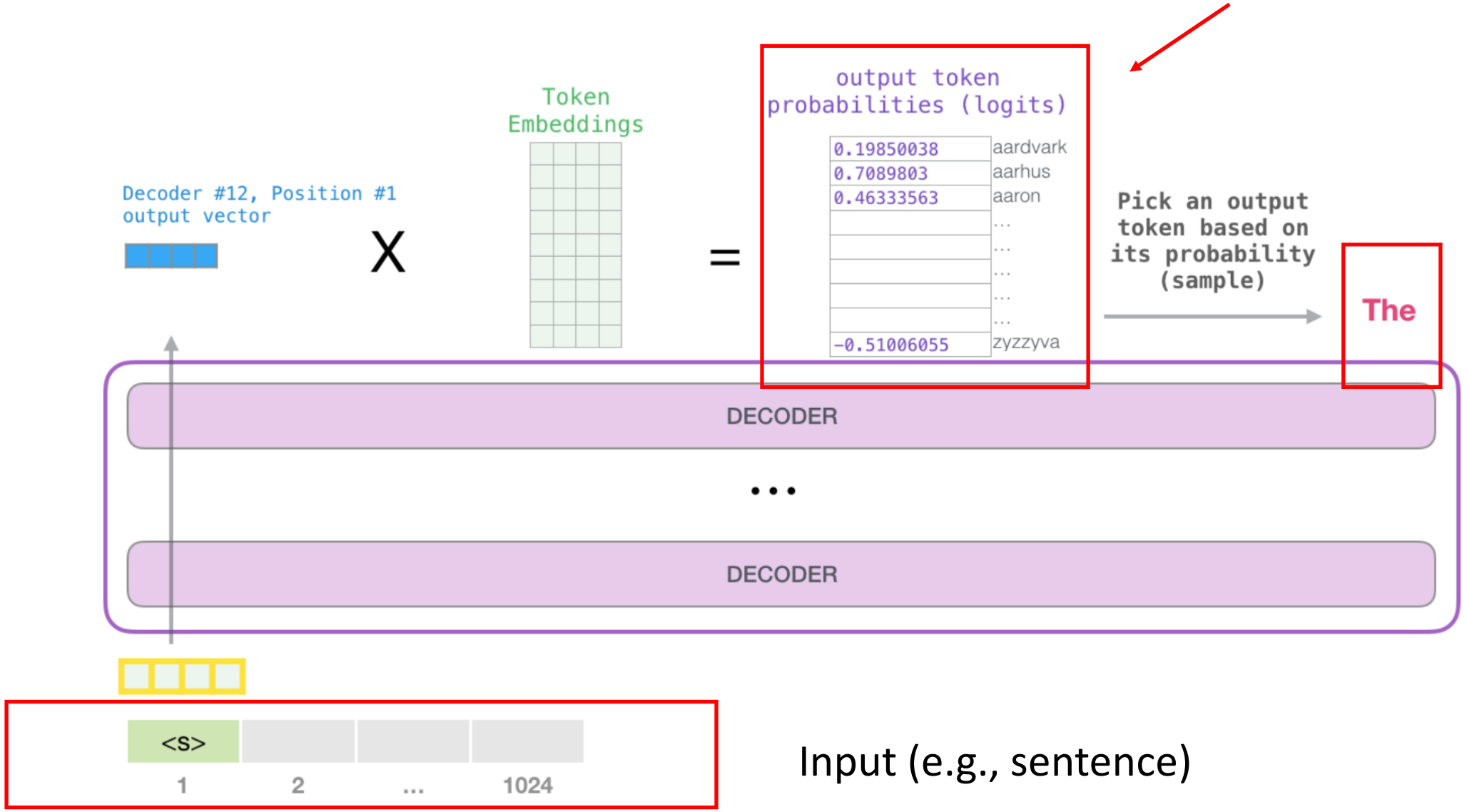
https://jalammar.github.io/illustrated-gpt2/



https://jalammar.github.io/illustrated-gpt2/



https://jalammar.github.io/illustrated-gpt2/



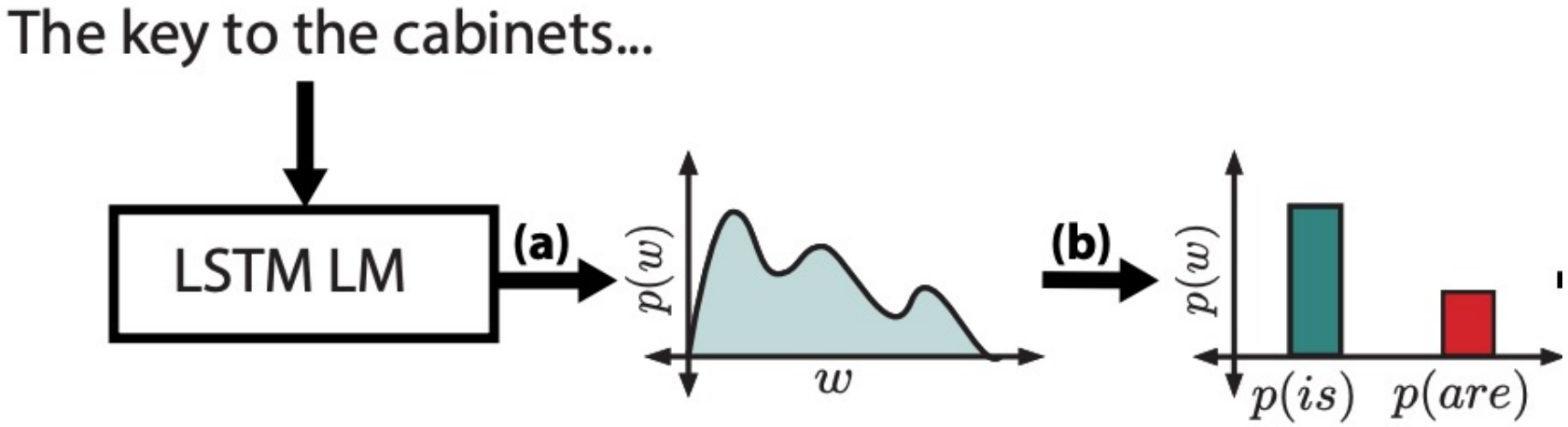
Input (e.g., sentence)

https://jalammar.github.io/illustrated-gpt2/

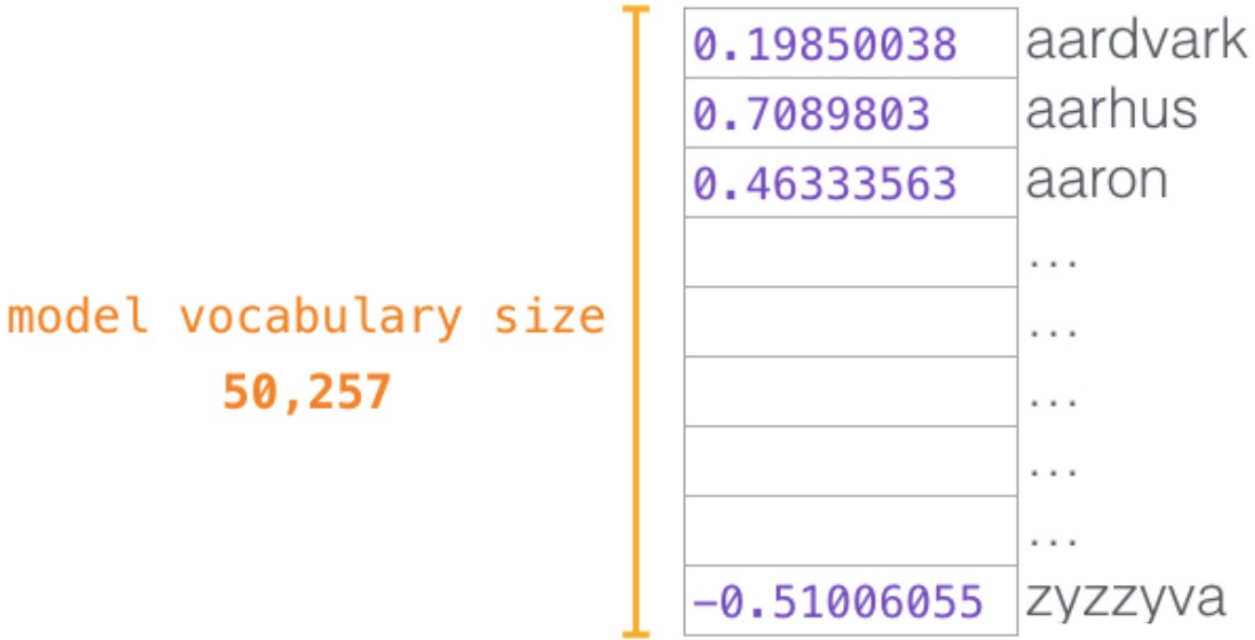
Key notions

- **Model output**: probability over words given prior context.
    - Given context “The key to the cabinet …”, how likely is it see word_1, word_2, word3, … word_n?
    - Assigning probability over all “words” in the model vocabularly.

Arehalli & Linzen (2020)



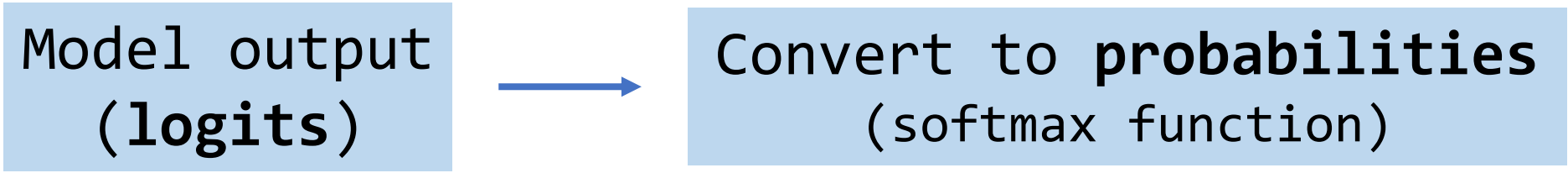
output token probabilities (logits)

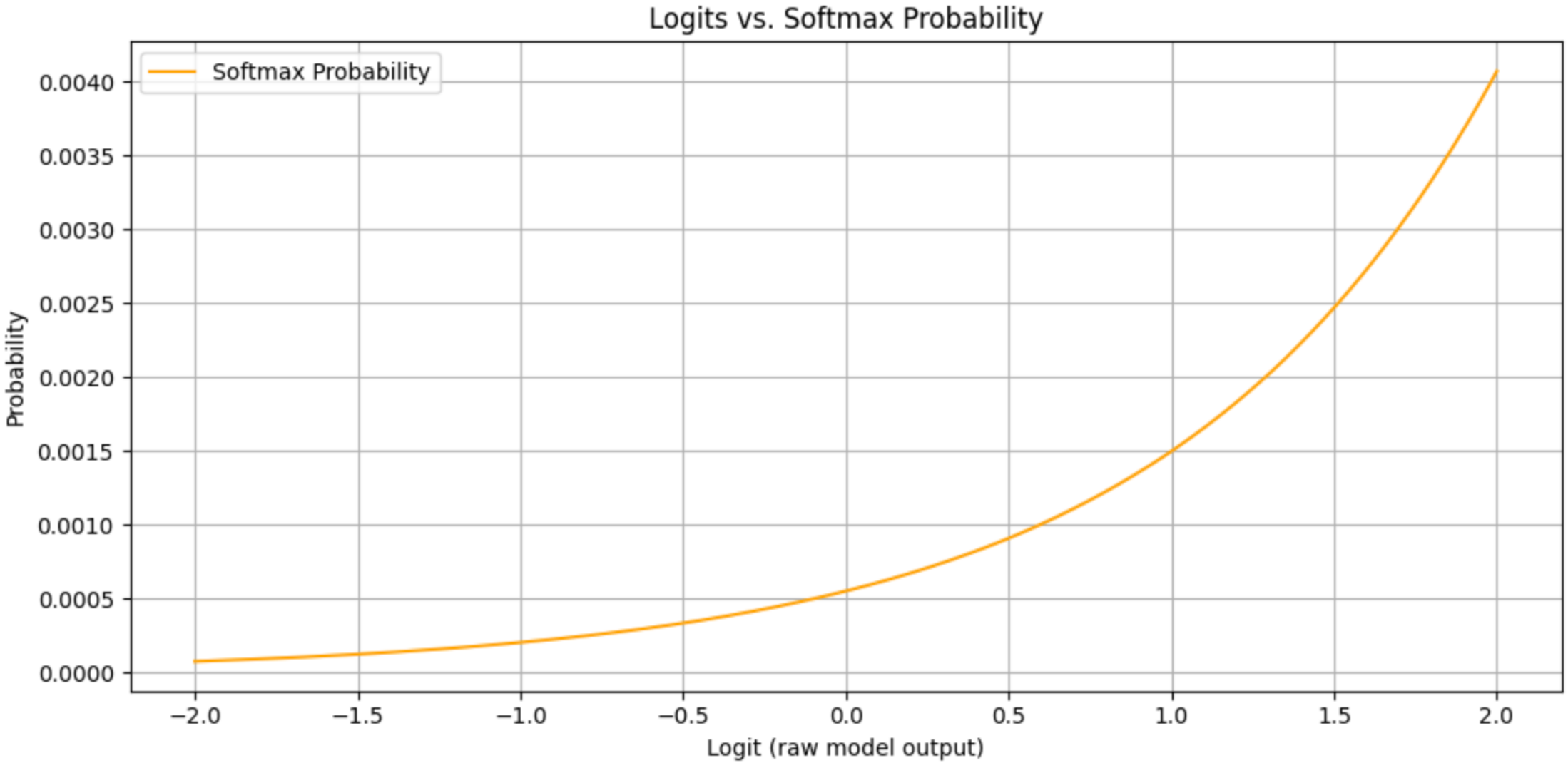


We will convert this to **log probability** and then **surprisal**.

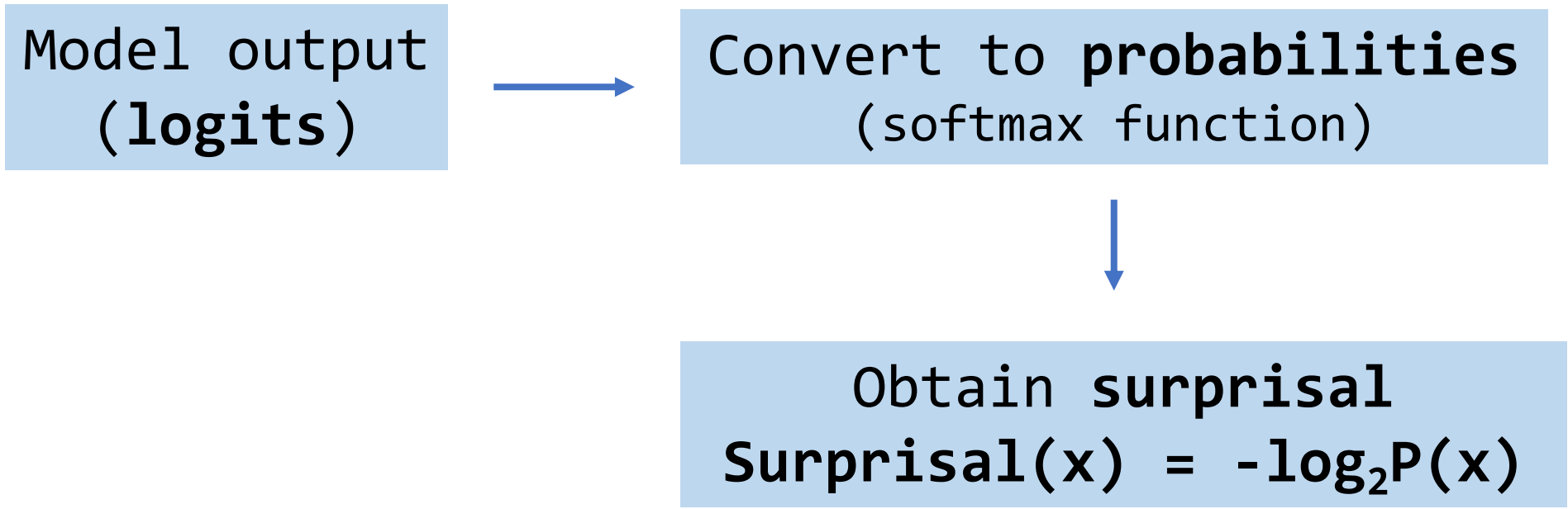
https://jalammar.github.io/illustrated-gpt2/

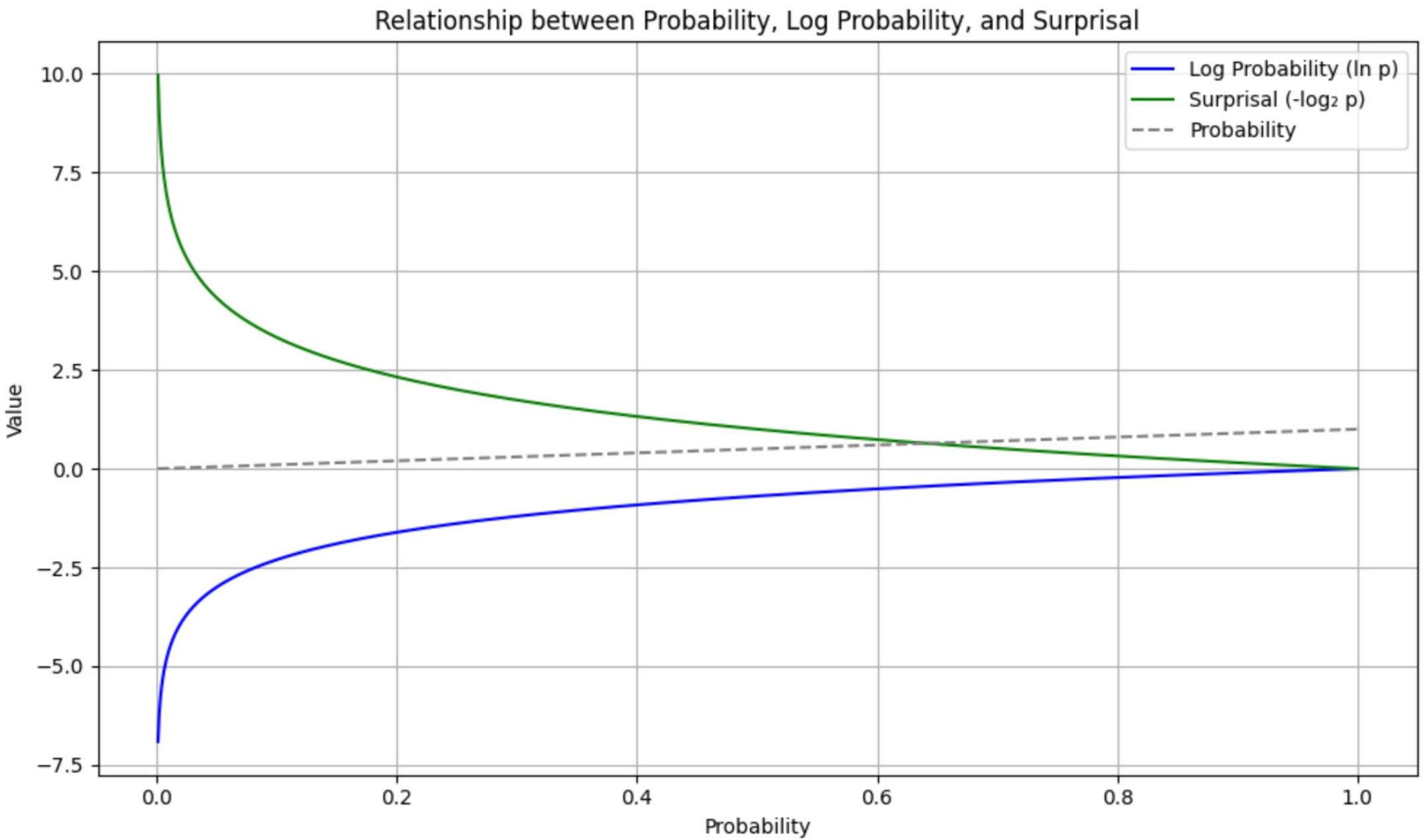
Workflow


Model output
(**logits**)
Convert to **probabilities**
(softmax function)

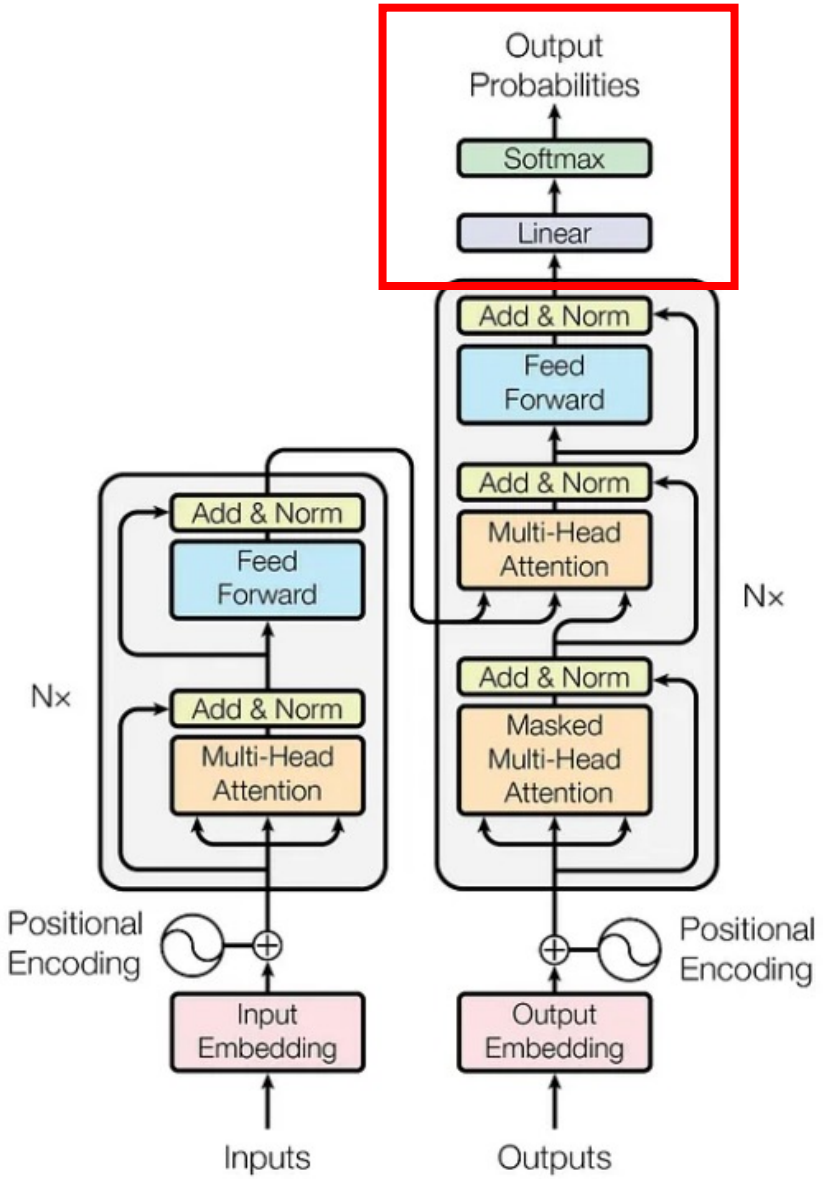

Logits vs. Softmax Probability
Softmax Probability
Probability
0.0040
0.0035
0.0030
0.0025
0.0020
0.0015
0.0010
0.0005
0.0000
−2.0
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
Logit (raw model output)

Workflow


Model output
(**logits**)
Convert to **probabilities**
(softmax function)
Obtain **surprisal**
Surprisal(x) = -log$_2$P(x)


Relationship between Probability, Log Probability, and Surprisal
Log Probability (ln p)
Surprisal (-log₂ p)
Probability
Value
Probability
10.0
7.5
5.0
2.5
0.0
−2.5
−5.0
−7.5
0.0
0.2
0.4
0.6
0.8
1.0

- Obtain model surprisal at the critical word
- Compare it with human reading time results



Transformer Architecture

Demo

- https://colab.research.google.com/drive/1u-AHUUWfCIX4WyB6t23NY_SlRPFnhpjt?usp=sharing