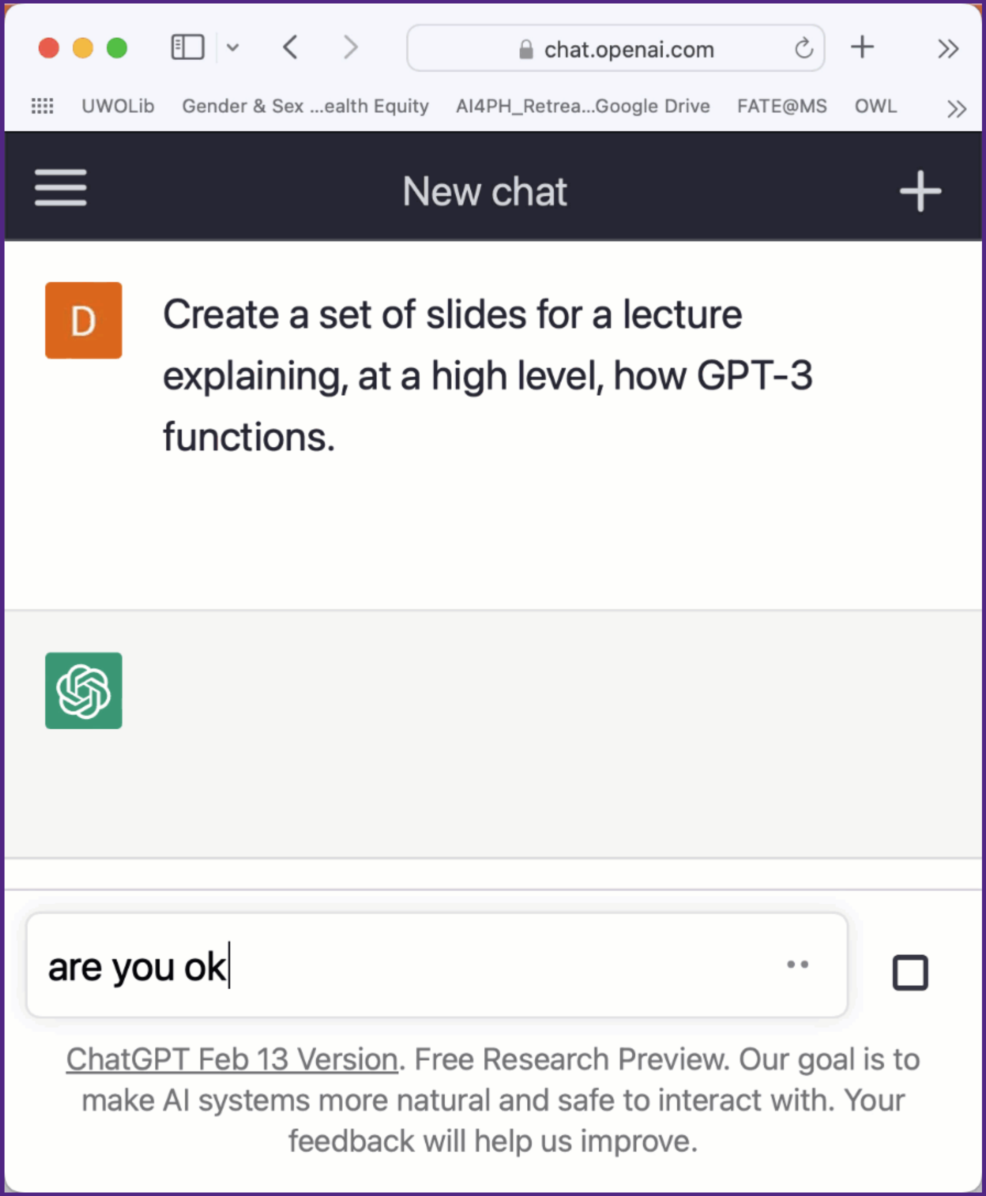
Neural Network Language Models

CS 4417B

The University of Western Ontario

GPT-x


chat.openai.com
UWOLib
Gender & Sex ...ealth Equity
AI4PH_Retrea...Google Drive
FATE@MS
OWL
New chat
D
Create a set of slides for a lecture explaining, at a high level, how GPT-3 functions.
are you ok
ChatGPT Feb 13 Version. Free Research Preview. Our goal is to make AI systems more natural and safe to interact with. Your feedback will help us improve.

Neural Network Language Models

- Input: Sequence of words
- Output: Probability over next word

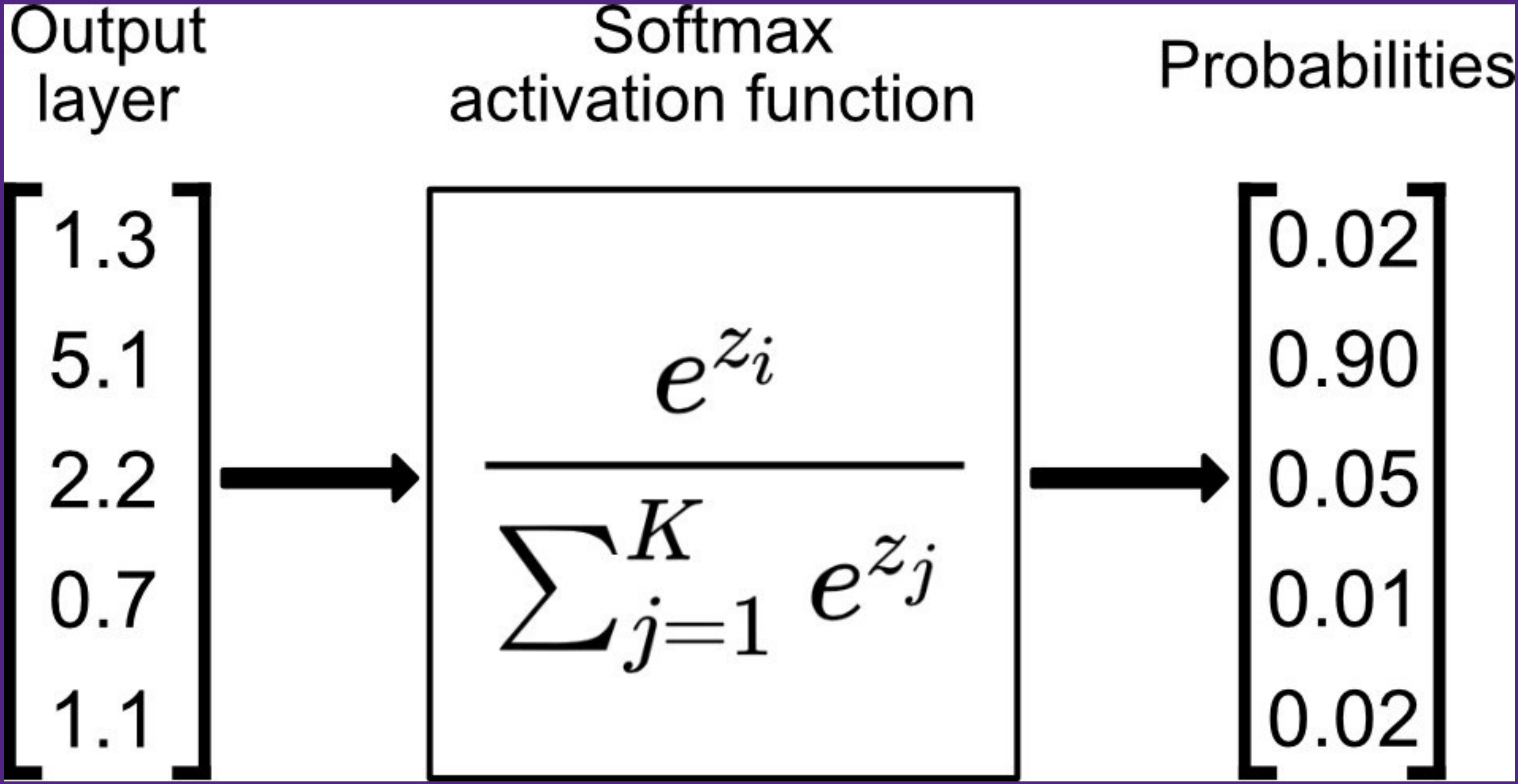
- Instead of storing a number for every *n*-gram, distribution is computed using the weights in the network.

- I am just giving a **high level overview** of how things work. Best detailed [NOT REQUIRED WATCHING] video is https://youtu.be/kCc8FmEb1nY
(Requires neural network background.)

Ideas in GPT

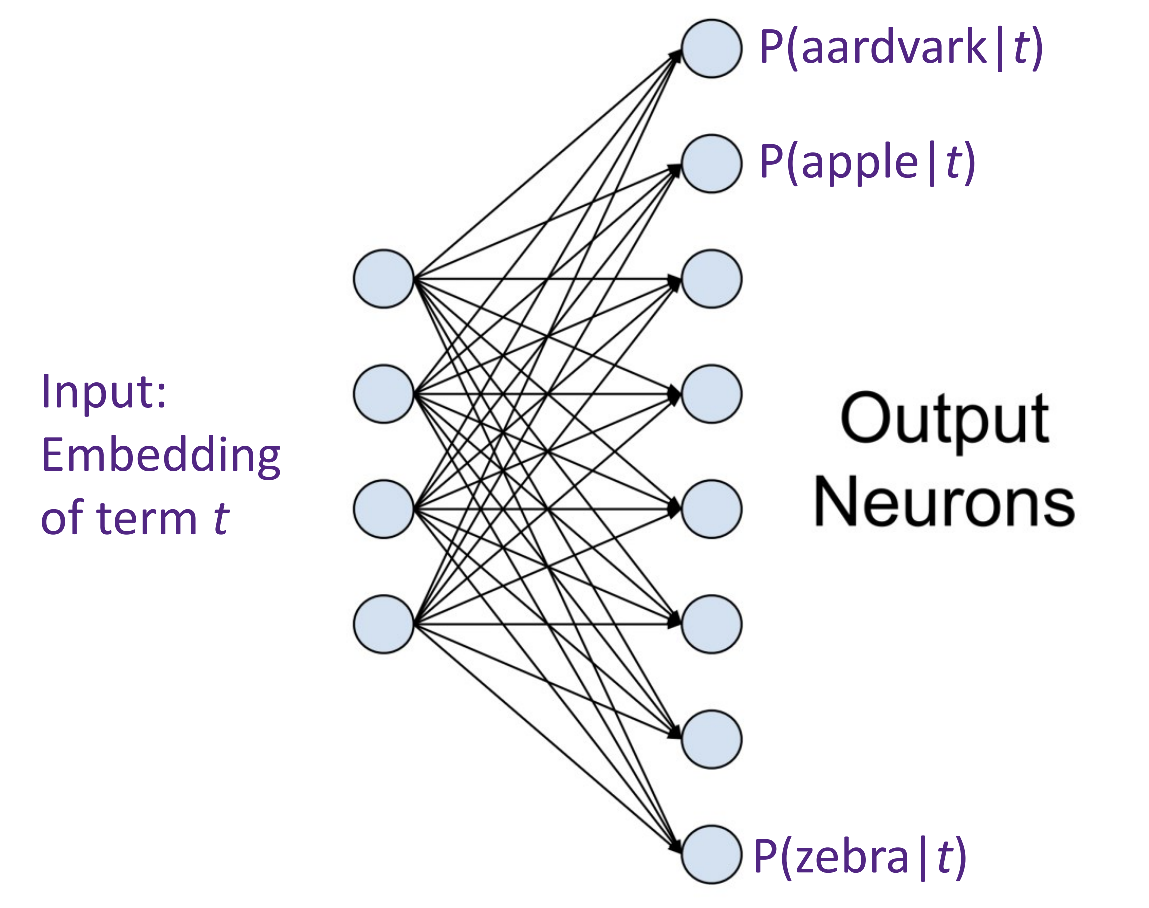
- Softmax output
- Hidden layers
- “Attention” to communicate information from word to word (big deal)
- Word embeddings and position information
- Byte-pair encodings

Reminder: Softmax



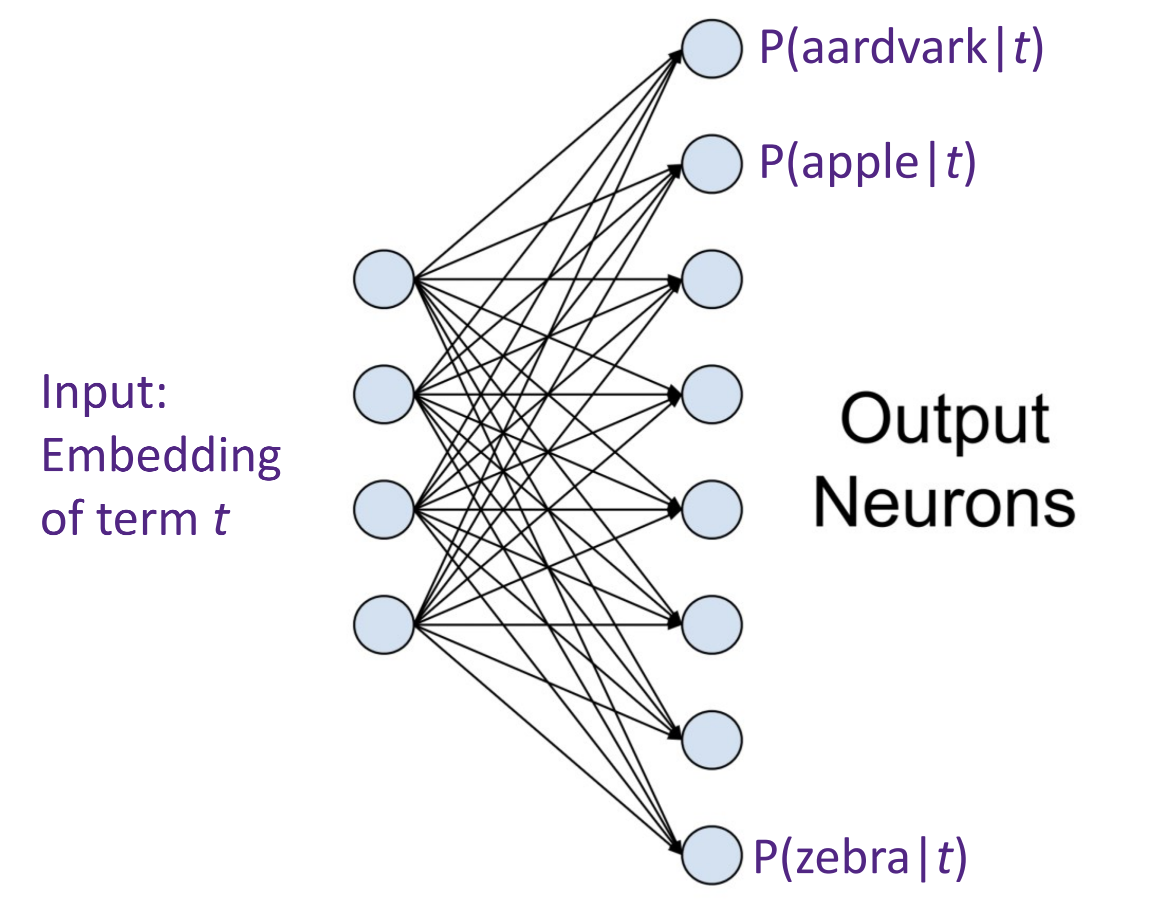
Output layer

- Produces probability distribution over tokens



Output layer

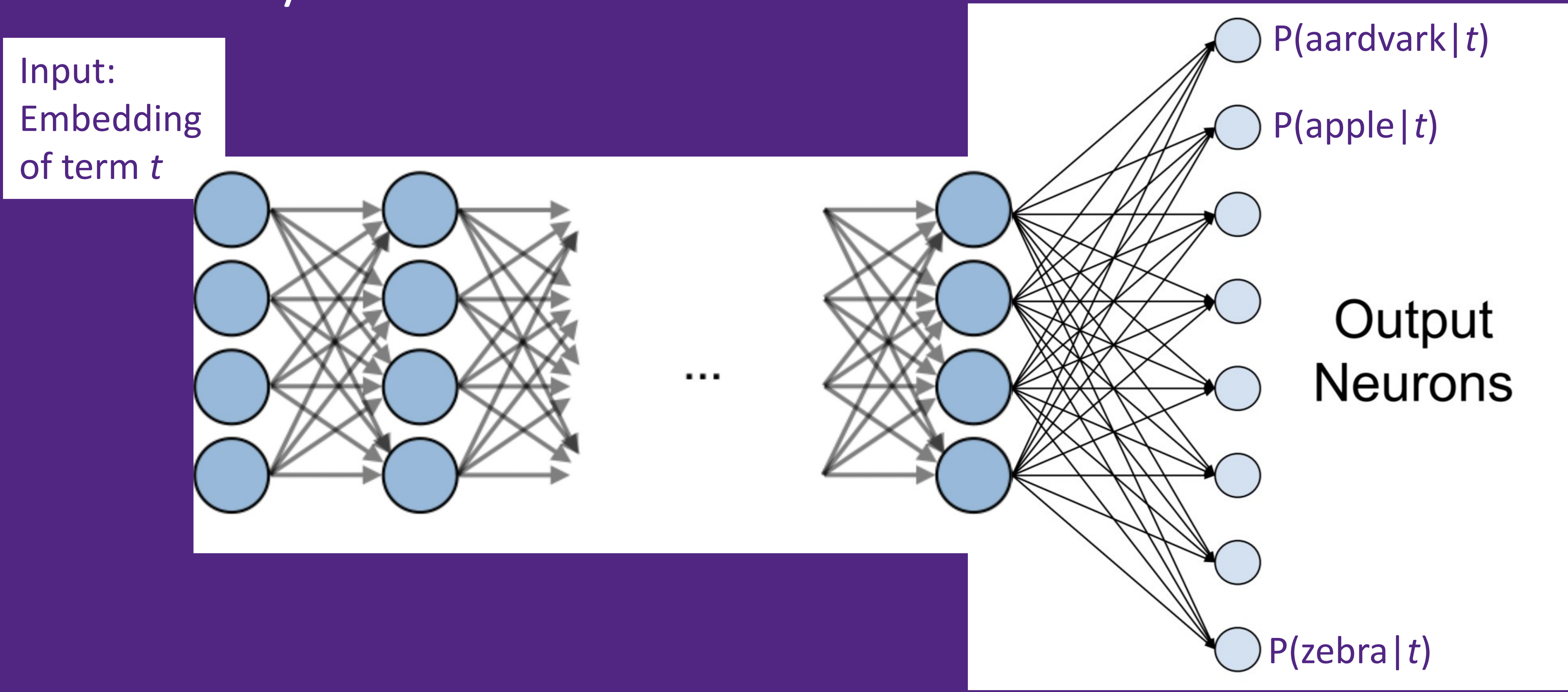
- Produces probability distribution over tokens



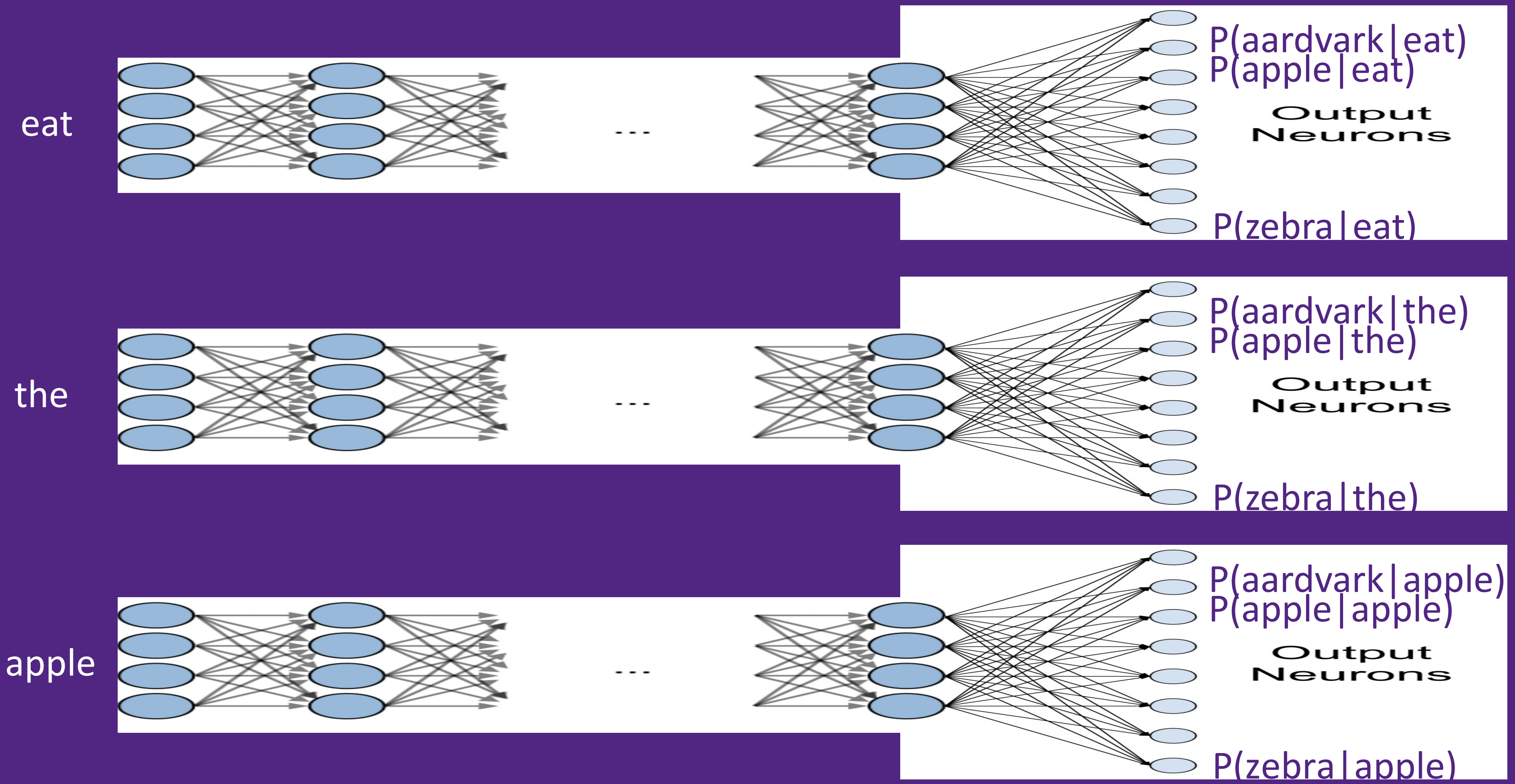
A) m*(m-1)   B) m*m   C) m*p   D) p*(m-1)

Output layer after hidden layers

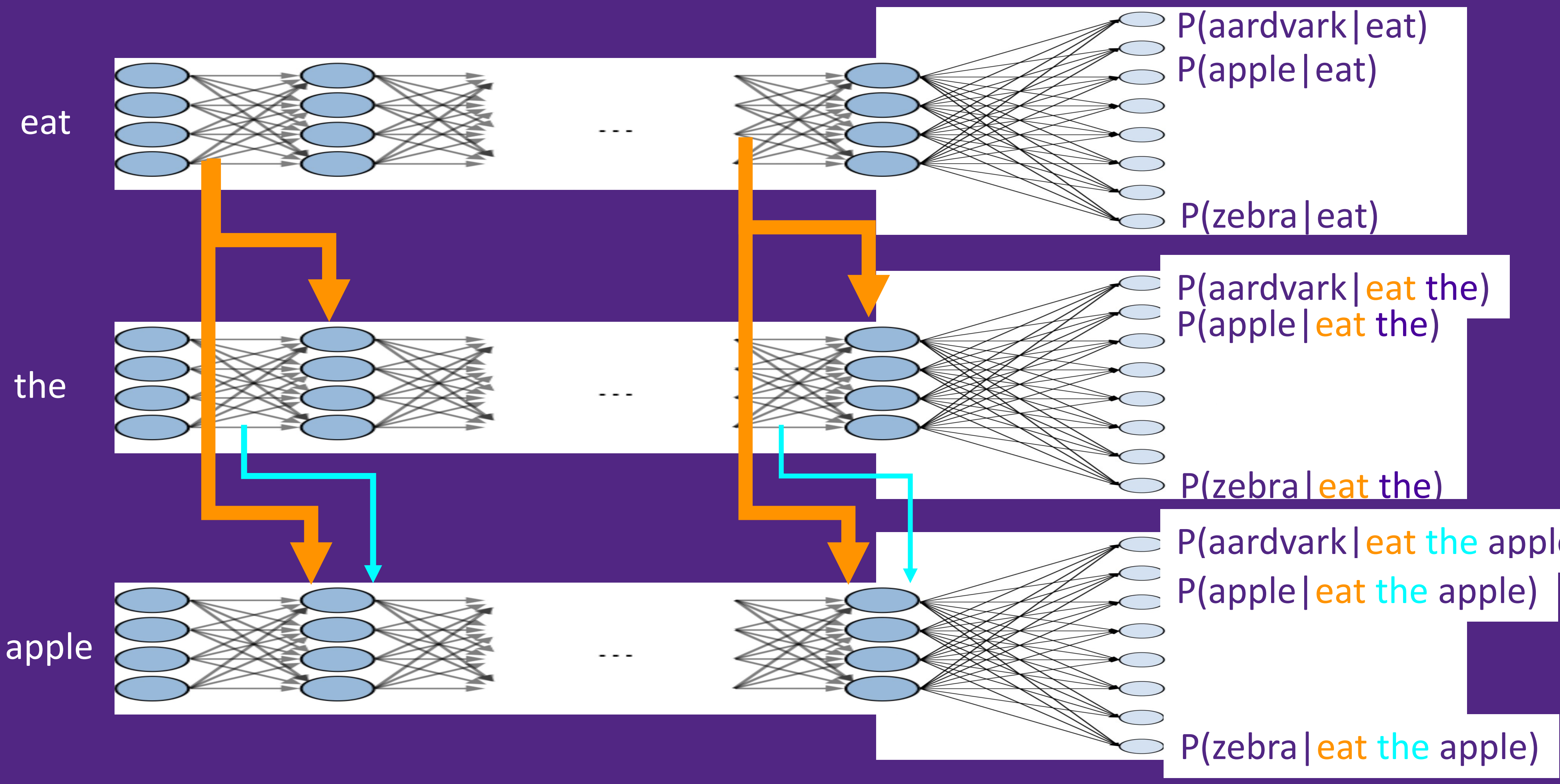
- Produces probability distribution over terms/tokens



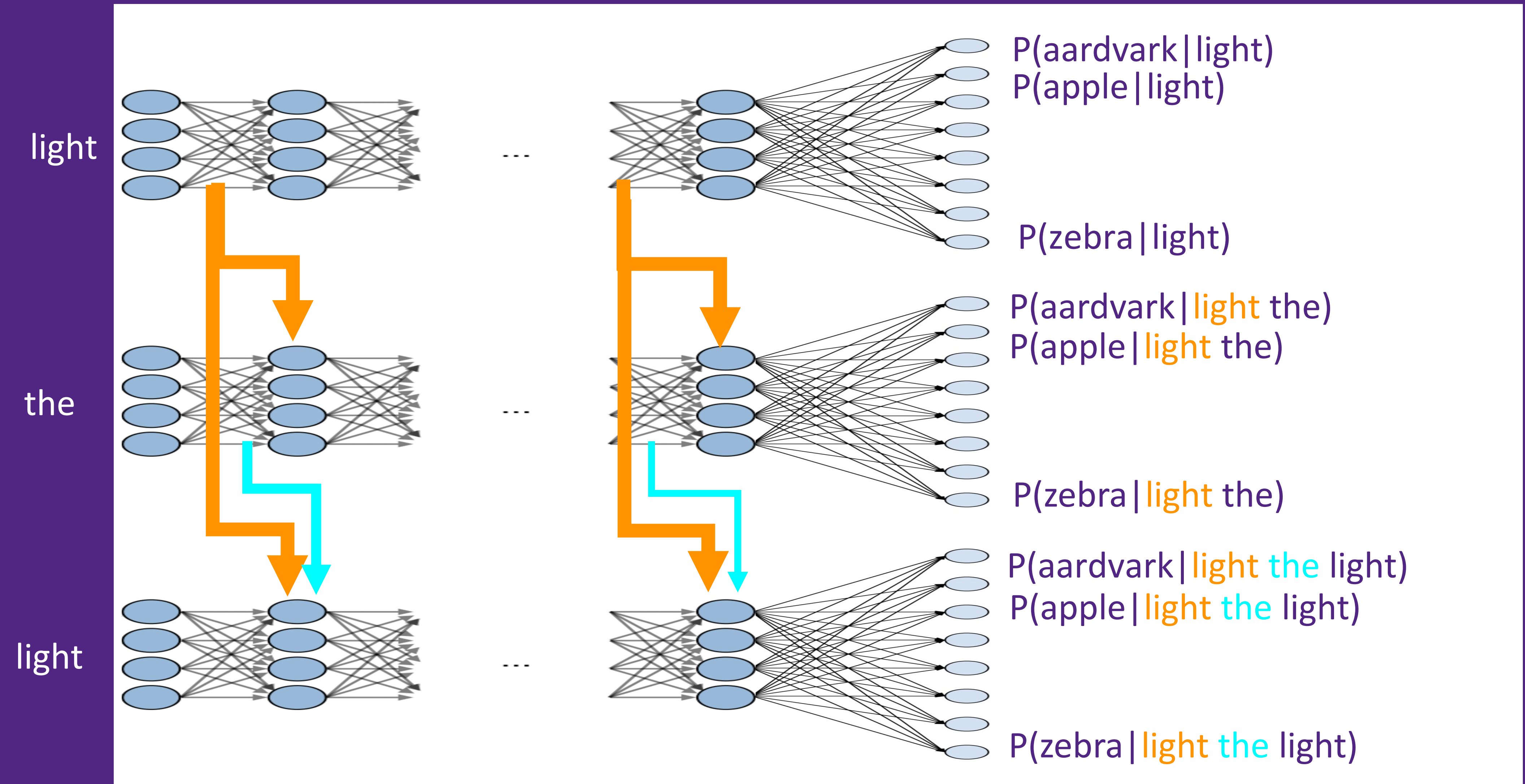
Simultaneous computations



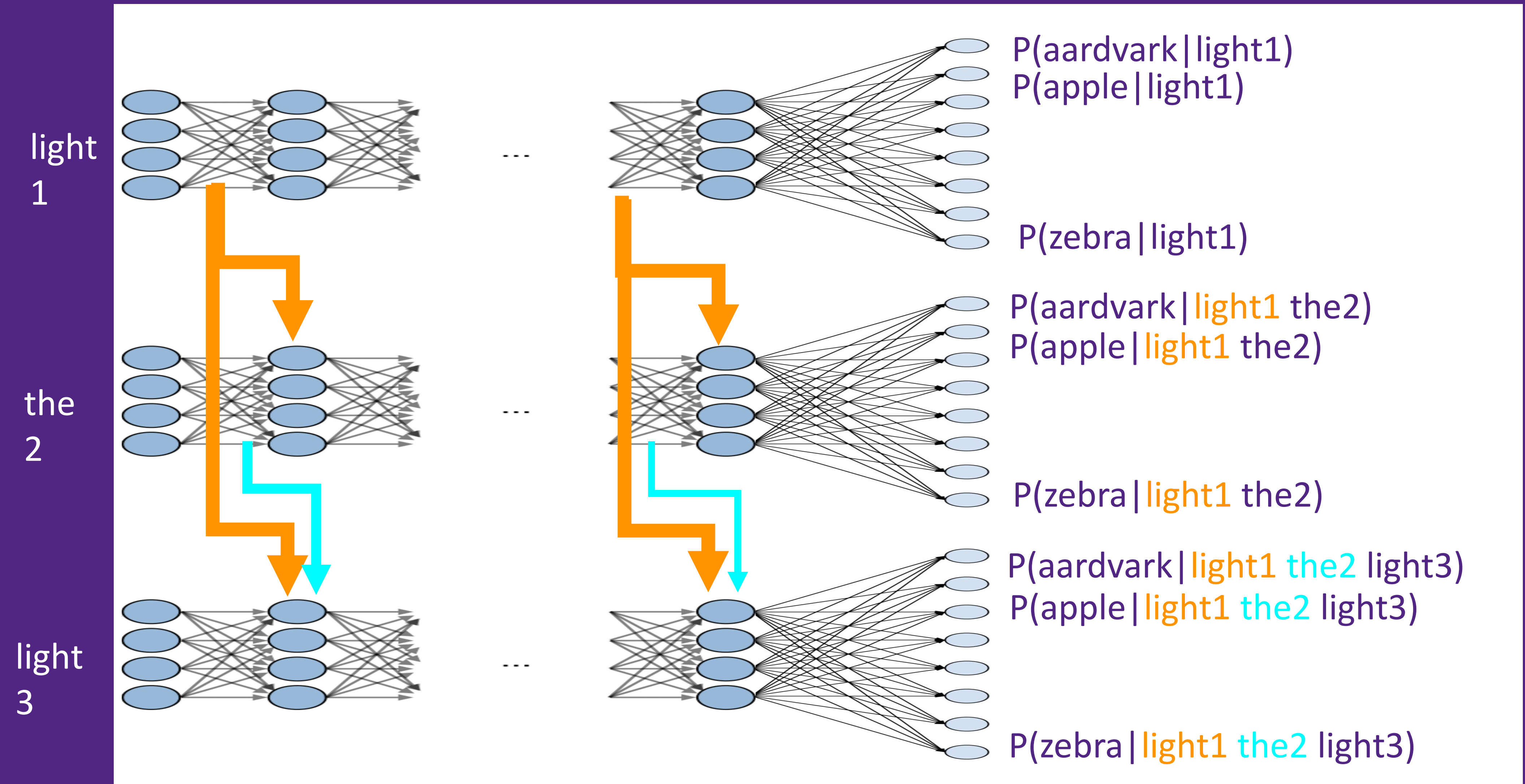
“Attention”


eat
the
apple
P(aardvark|eat)
P(apple|eat)
P(zebra|eat)
P(aardvark|eat the)
P(apple|eat the)
P(zebra|eat the)
P(aardvark|eat the appl
P(apple|eat the apple)
P(zebra|eat the apple)

Positional Encodings



Positional Encodings



Tokenization:
“Byte-Pair Encoding”

- Issue: “Out-of vocabulary” words

- Idea
  - Fix number of allowed tokens
  - Allocate tokens to frequently occurring sequences of characters
  - Still allow any sequence of characters to be mapped to tokens

- GPT-2: 256 one-byte tokens + 50000 learned tokens + 1 special end-of-sequence token

Byte pair encoding

- Start with 1 token per byte
- If a pair of bytes occurs frequently,
  make a new token
  - If “t” “h” occurs a lot, create a new token “t” “h” and a merge rule.
  - If “th” “e” occurs a lot, create a new token “the” and a merge rule.
  - Stop when bored. (For GPT, after 50000 merges.)

- To tokenize a new sentence, apply the rules in order.
- Common sequences will get their own token; uncommon ones won’t
- *Any* sequence of bytes is tokenizable
- In GPT, one token is about ¾ of a word, on average.
- https://platform.openai.com/tokenizer

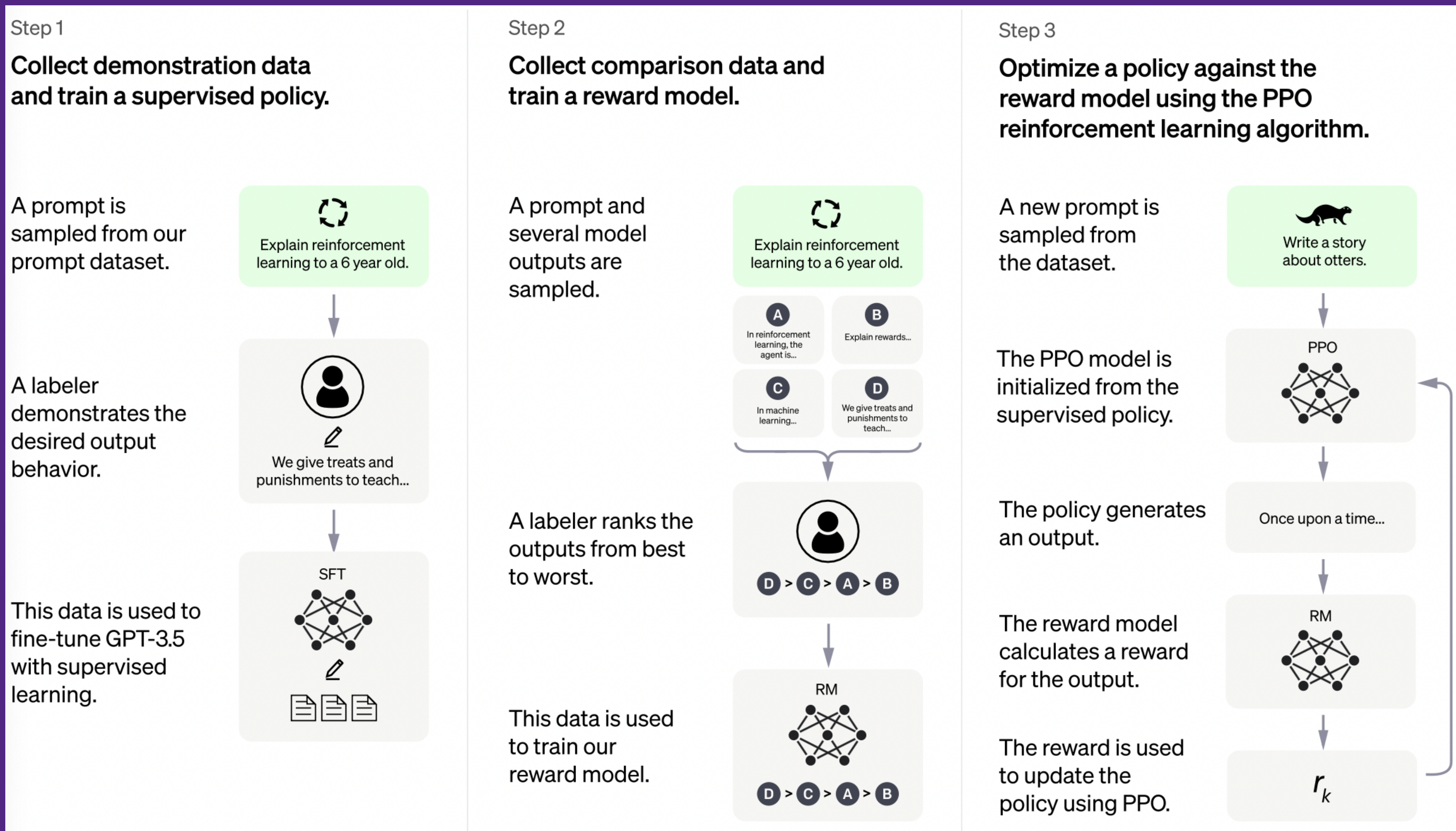
Parameters

- GPT-2
  - Vocabulary size: 50,257
  - Context length 1024
  - 10^9.18 parameters
- GPT-3
  - Vocabulary size: 50,257
  - Context length 2048
  - 10^11.25 parameters

Parameters

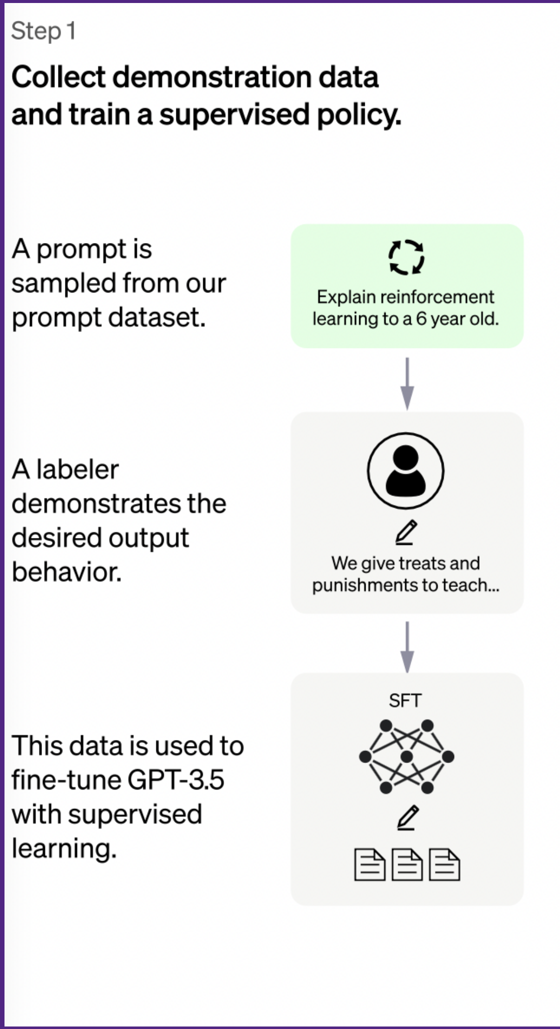
- GPT-2
  - Vocabulary size: 50,257
  - Context length 1024
  - 10^9.18 parameters
  - Count-based: ~10^4814 parameters
- GPT-3 (English part)
  - Vocabulary size: 50,257
  - Context length 2048
  - 10^11.25 parameters
  - Count-based ~10^9628 parameters

ChatGPT



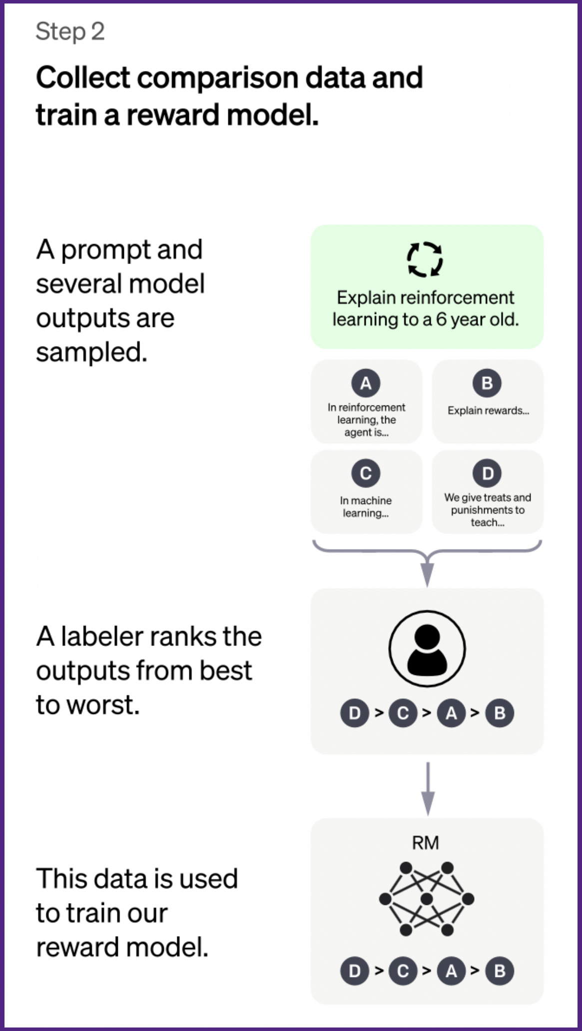
Step 1
Collect demonstration data and train a supervised policy.
A prompt is sampled from our prompt dataset.
Explain reinforcement learning to a 6 year old.
A labeler demonstrates the desired output behavior.
We give treats and punishments to teach...
This data is used to fine-tune GPT-3.5 with supervised learning.
SFT
Step 2
Collect comparison data and train a reward model.
A prompt and several model outputs are sampled.
Explain reinforcement learning to a 6 year old.
A
In reinforcement learning, the agent is...
B
Explain rewards...
C
In machine learning...
D
We give treats and punishments to teach...
A labeler ranks the outputs from best to worst.
D > C > A > B
This data is used to train our reward model.
RM
D > C > A > B
Step 3
Optimize a policy against the reward model using the PPO reinforcement learning algorithm.
A new prompt is sampled from the dataset.
Write a story about otters.
The PPO model is initialized from the supervised policy.
PPO
The policy generates an output.
Once upon a time...
The reward model calculates a reward for the output.
RM
The reward is used to update the policy using PPO.
r_k

Fine-tune GPT for prompt/response



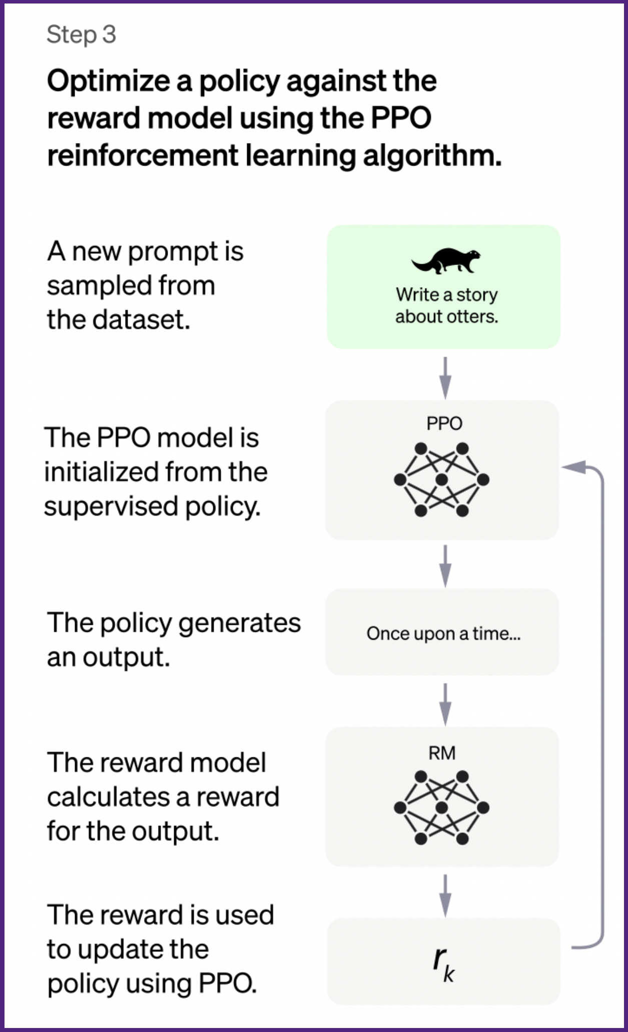
- Use *human-created* pairs of prompts and responses
- Fine-tune the language production model

Learn a model that can rank responses



- Use human-derived *rankings* of responses
- Training pairs are (model outputs, rankings)
- Learn to predict the ranking of a series of outputs (which is best?)

Train a more sophisticated babbling model based on the reward model



- Learn how to *produce* good responses by trying to produce high rewards according to the learned reward model
- Human not required at this point.
- https://arxiv.org/pdf/2203.02155.pdf

Concerns about Large Language Models

- On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🦜

- FAccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency
- March 2021
- Pages 610–623
- https://doi.org/10.1145/3442188.3445922

Environmental

- Average human is responsible for an estimated 5t $CO2$

- Authors trained a Transformer (big) model [136] with neural architecture search and estimated that the training procedure emitted 284t of $CO2$.

Representation

- Pew Internet Research's 2016 survey reveals 67% of Reddit users in the United States are men, and 64% between ages 18 and 29.13
- Similarly, recent surveys of Wikipedians find that only 8.8–15% are women or girls [9].
- The training set for GPT-3 was a filtered version of the Common Crawl dataset, developed by training a classifier to pick out those documents most similar to the ones used in GPT-2's training data, i.e. documents linked to from Reddit [25], plus Wikipedia and a collection of books.

Data changes – what is updated?

- For instance, the Black Lives Matter movement (BLM) influenced Wikipedia article generation and editing such that, as the BLM movement grew, articles covering shootings of Black people in-creased in coverage and were generated with reduced latency [135].

Curation, Documentation, Accountability

- “Feeding AI systems on the world’s beauty, ugliness, and cruelty, but expecting it to reflect only the beauty is a fantasy.” [p.1541]
- [15] Ruha Benjamin. 2019. Race After Technology: Abolitionist Tools for the New Jim Code. Polity Press, Cambridge, UK.

The Stochastic Parrot 🦜

- Language models repeat what they are taught, injecting controlled randomness to create new and different text.

- Because LMs produce text that is so convincing, it’s important to consider what went into their construction when assessing the credibility of what they produce.

He walked on the moon & walked up Ararat, searching in vain for Noah's Ark.

I am not aware of any astronaut who has walked on the moon and also walked on Mount Ararat searching for Noah's Ark. The only lunar missions were carried out by NASA as part of the Apollo program between 1969 and 1972, and none of the astronauts on those missions were known to have searched for Noah's Ark on Mount Ararat. Mount Ararat is a mountain located in Turkey, near the borders of Armenia and Iran, that is said in some religious texts to be the final resting place of Noah's Ark after the biblical flood. However, there is no scientific evidence to support the existence of a historical ark, or that a flood as described in the Bible ever occurred.

https://en.wikipedia.org/wiki/James_Irwin