What is an LLM?

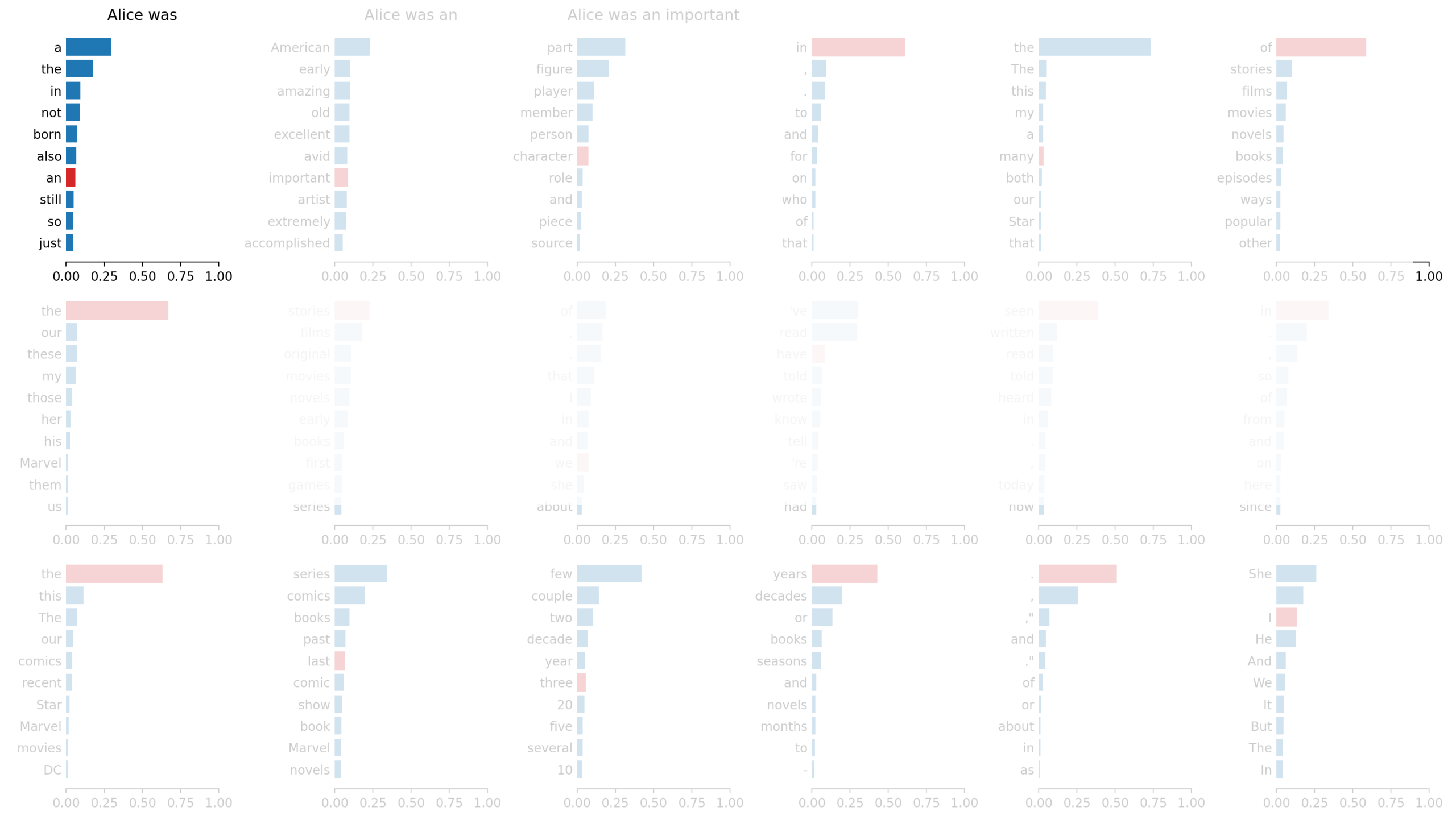
Large Language Models

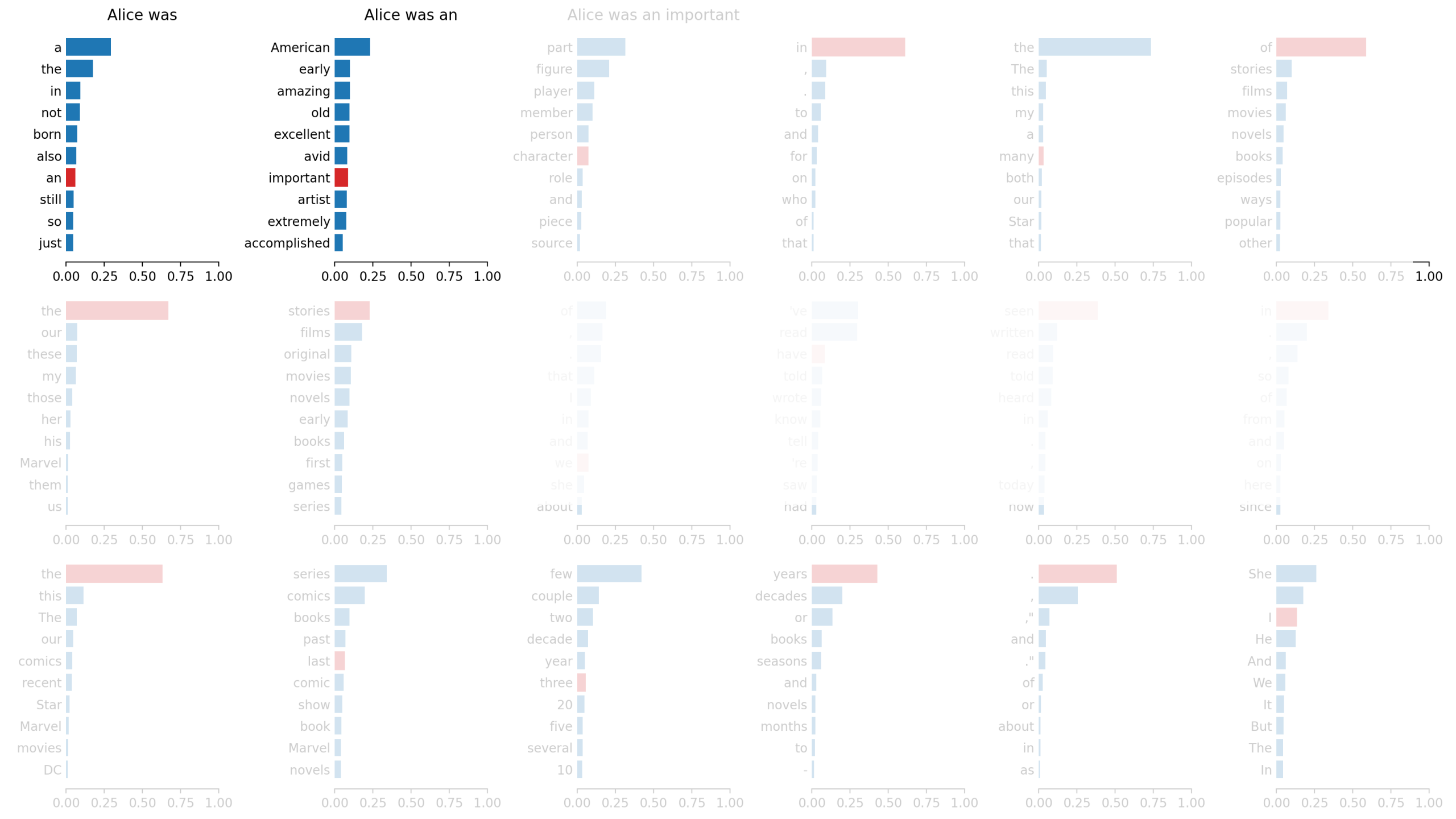
Predict the next word

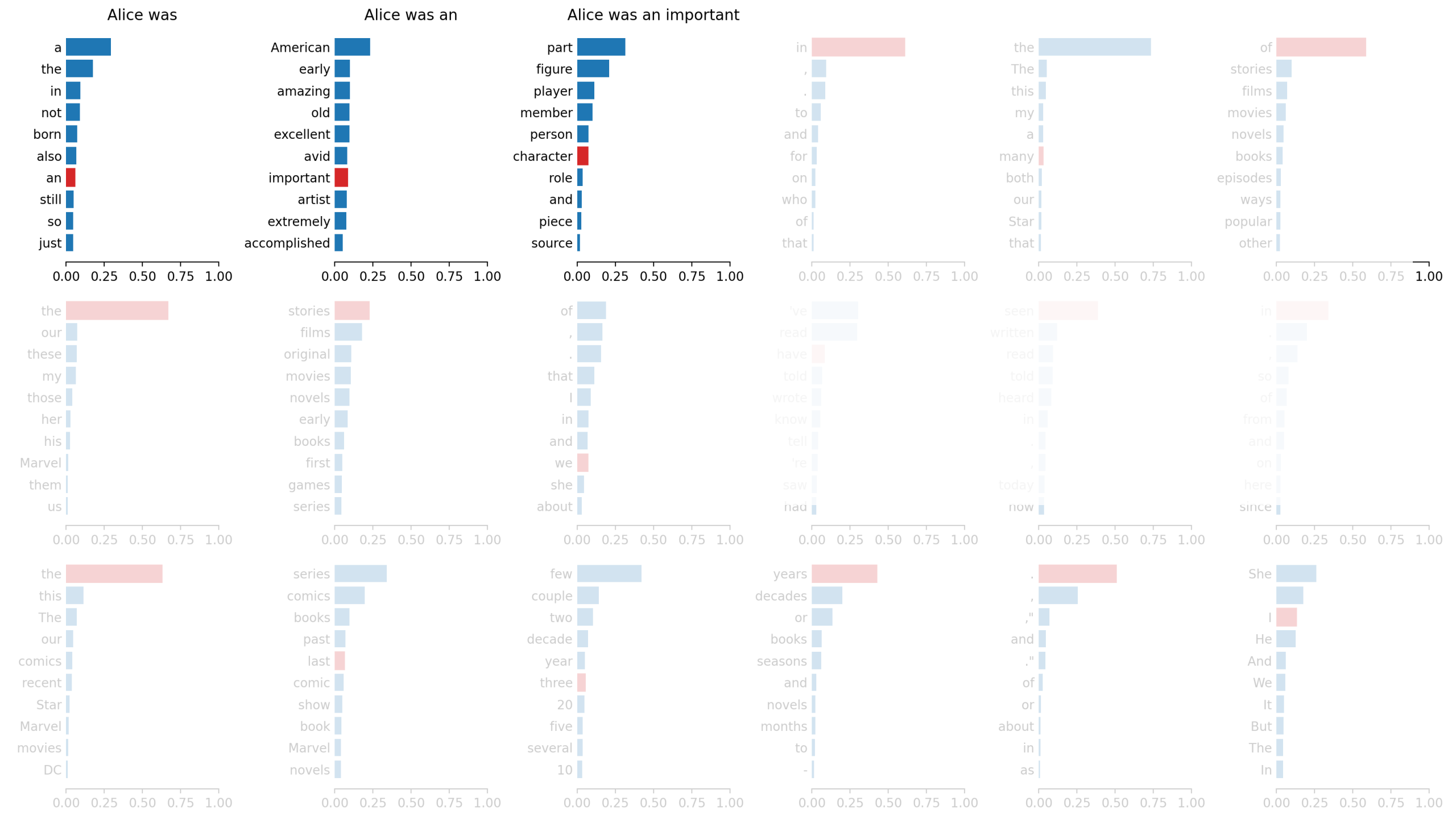
Alice was

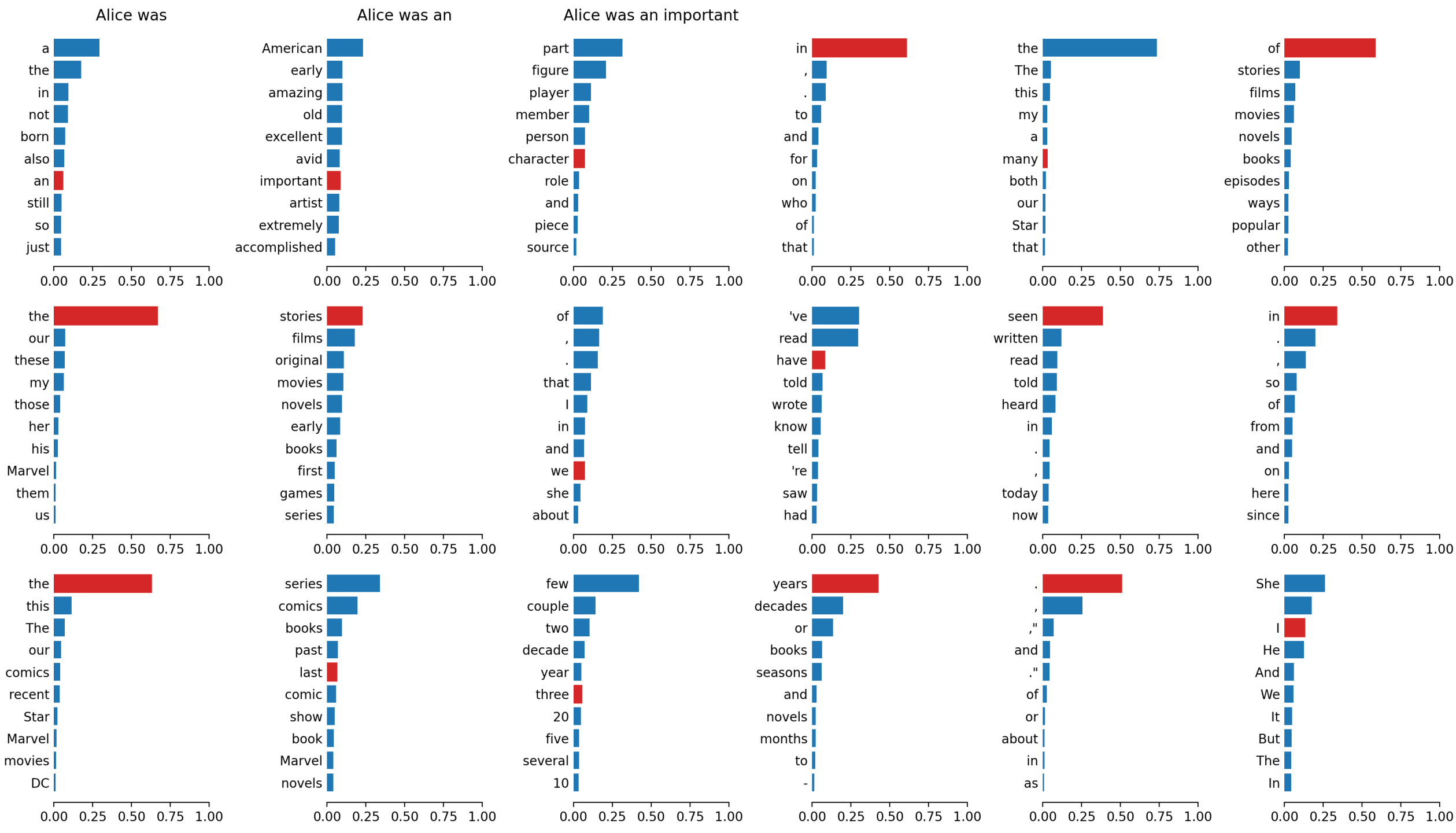
Predict the next word

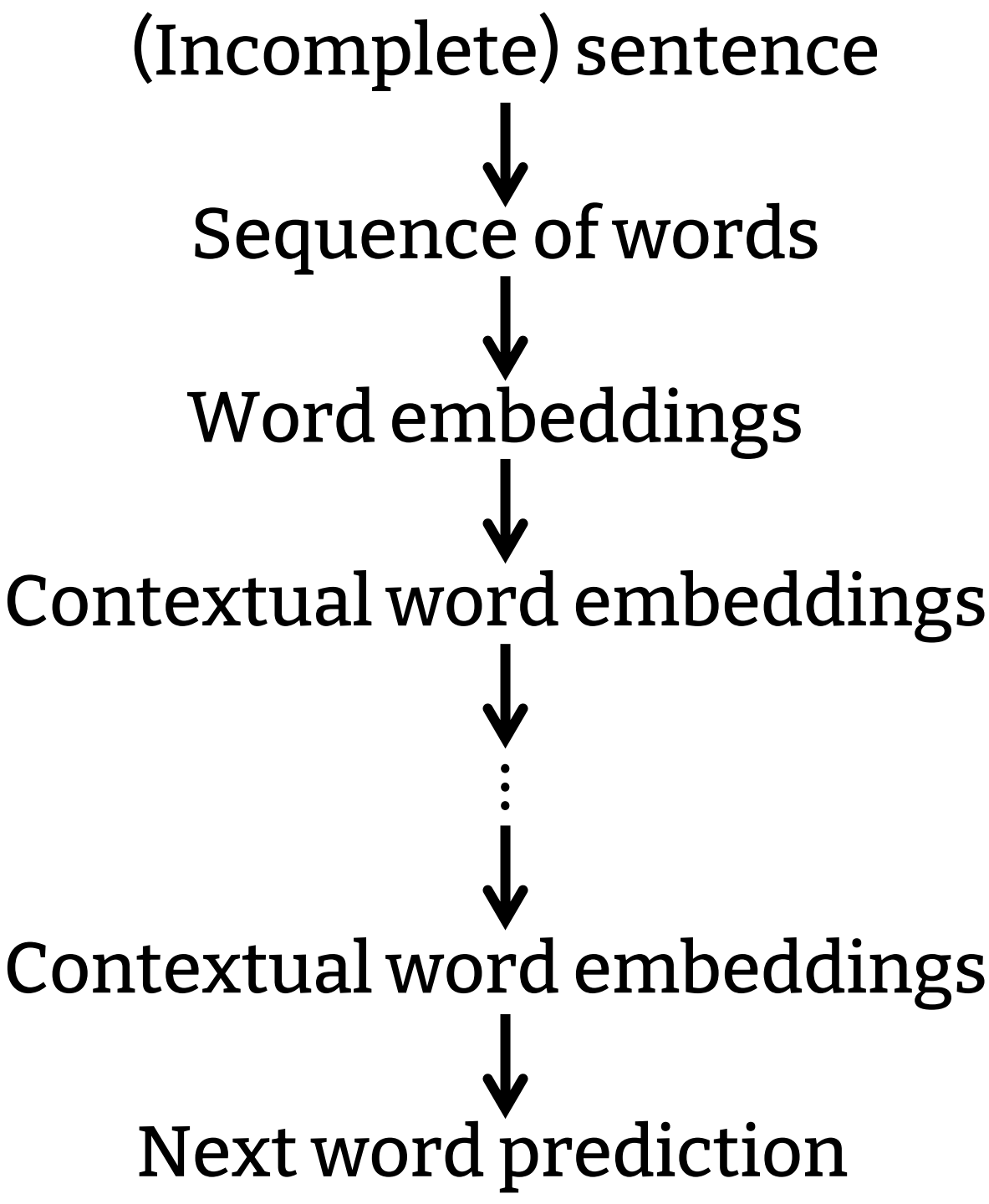
Alice was an important character in many of the stories we have seen in the last three years. I'm very proud to be able to bring back that character. And it's going to take an incredible amount of time to do that."
The film has been nominated for several Academy Awards, and is scheduled for a May 4, 2018 release.<|endoftext|>

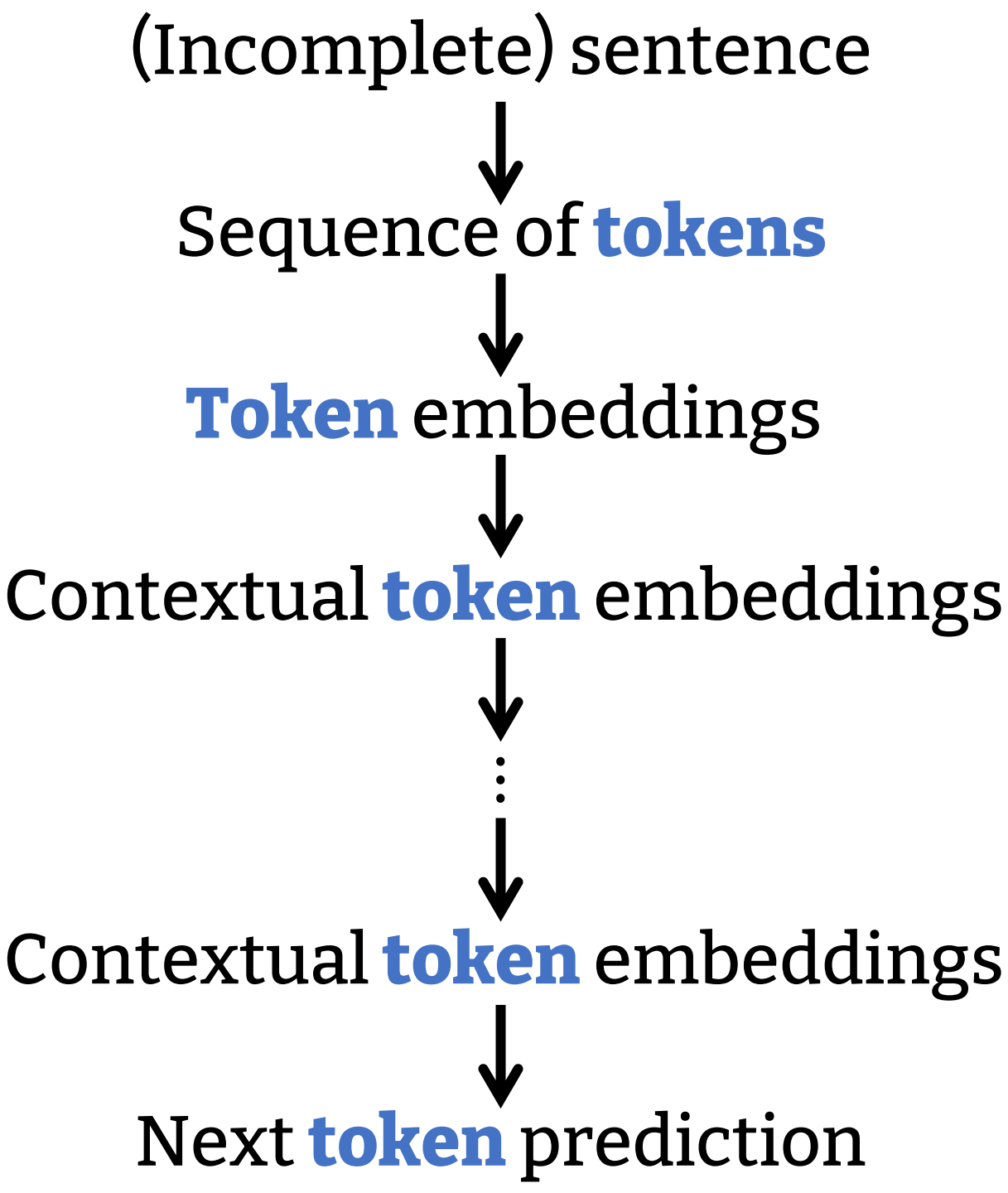

Alice was
a
the
in
not
born
also
an
still
so
just
Alice was an
American
early
amazing
old
excellent
avid
important
artist
extremely
accomplished
Alice was an important
part
figure
player
member
person
character
role
and
piece
source
in
,
.
to
and
for
on
who
of
that
the
The
this
my
a
many
both
our
Star
that
of
stories
films
movies
novels
books
episodes
ways
popular
other
the
our
these
my
those
her
his
Marvel
them
us
stories
films
original
movies
novels
early
books
first
games
series
of
,
.
that
I
in
and
we
she
about
've
read
have
told
wrote
know
tell
're
saw
had
seen
written
read
told
heard
in
.
,
today
now
in
.
,
so
of
from
and
on
here
since
0.00 0.25 0.50 0.75 1.00
the
this
The
our
comics
recent
Star
Marvel
movies
DC
series
comics
books
past
last
comic
show
book
Marvel
novels
few
couple
two
decade
year
three
20
five
several
10
years
decades
or
books
seasons
and
novels
months
to
-
.
,
,"
and
."
of
or
about
in
as
She
I
He
And
We
It
But
The
In

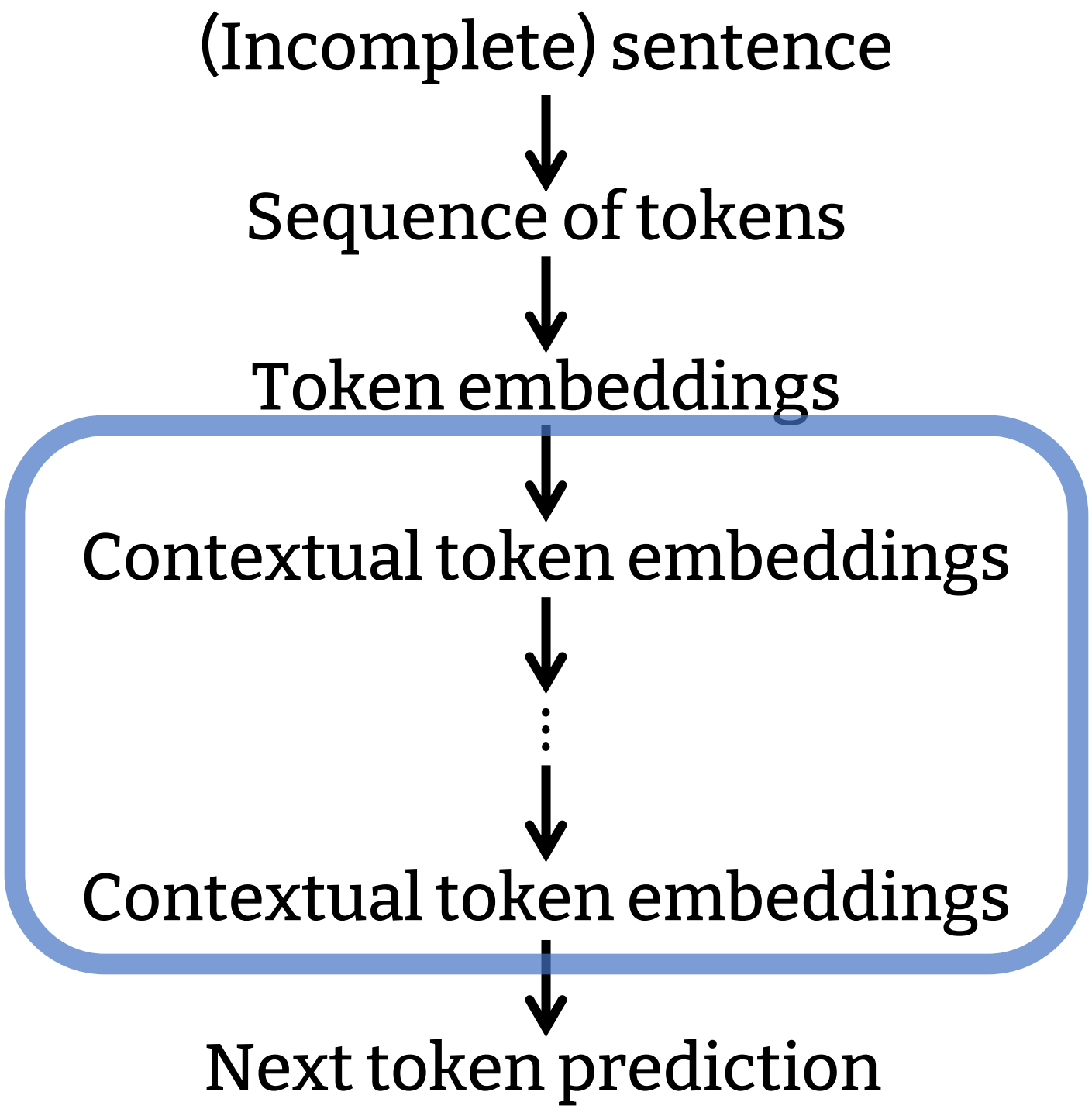

Alice was
a
the
in
not
born
also
an
still
so
just
Alice was an
American
early
amazing
old
excellent
avid
important
artist
extremely
accomplished
Alice was an important
part
figure
player
member
person
character
role
and
piece
source
in
,
.
to
and
for
on
who
of
that
the
The
this
my
a
many
both
our
Star
that
of
stories
films
movies
novels
books
episodes
ways
popular
other
the
our
these
my
those
her
his
Marvel
them
us
stories
films
original
movies
novels
early
books
first
games
series
of
,
.
that
I
in
and
we
she
about
've
read
have
told
wrote
know
tell
're
saw
had
seen
written
read
told
heard
in
.
,
today
now
in
.
,
so
of
from
and
on
here
since
the
this
The
our
comics
recent
Star
Marvel
movies
DC
series
comics
books
past
last
comic
show
book
Marvel
novels
few
couple
two
decade
year
three
20
five
several
10
years
decades
or
books
seasons
and
novels
months
to
-
.
,
,"
and
."
of
or
about
in
as
She
I
He
And
We
It
But
The
In
0.00
0.25
0.50
0.75
1.00


Alice was
a
the
in
not
born
also
an
still
so
just
Alice was an
American
early
amazing
old
excellent
avid
important
artist
extremely
accomplished
Alice was an important
part
figure
player
member
person
character
role
and
piece
source
in
,
.
to
and
for
on
who
of
that
the
The
this
my
a
many
both
our
Star
that
of
stories
films
movies
novels
books
episodes
ways
popular
other
the
our
these
my
those
her
his
Marvel
them
us
stories
films
original
movies
novels
early
books
first
games
series
of
,
.
that
I
in
and
we
she
about
've
read
have
told
wrote
know
tell
're
saw
had
seen
written
read
told
heard
in
.
,
today
now
in
.
,
so
of
from
and
on
here
since
the
this
The
our
comics
recent
Star
Marvel
movies
DC
series
comics
books
past
last
comic
show
book
Marvel
novels
few
couple
two
decade
year
three
20
five
several
10
years
decades
or
books
seasons
and
novels
months
to
-
.
,
,"
and
."
of
or
about
in
as
She
I
He
And
We
It
But
The
In
0.00
0.25
0.50
0.75
1.00

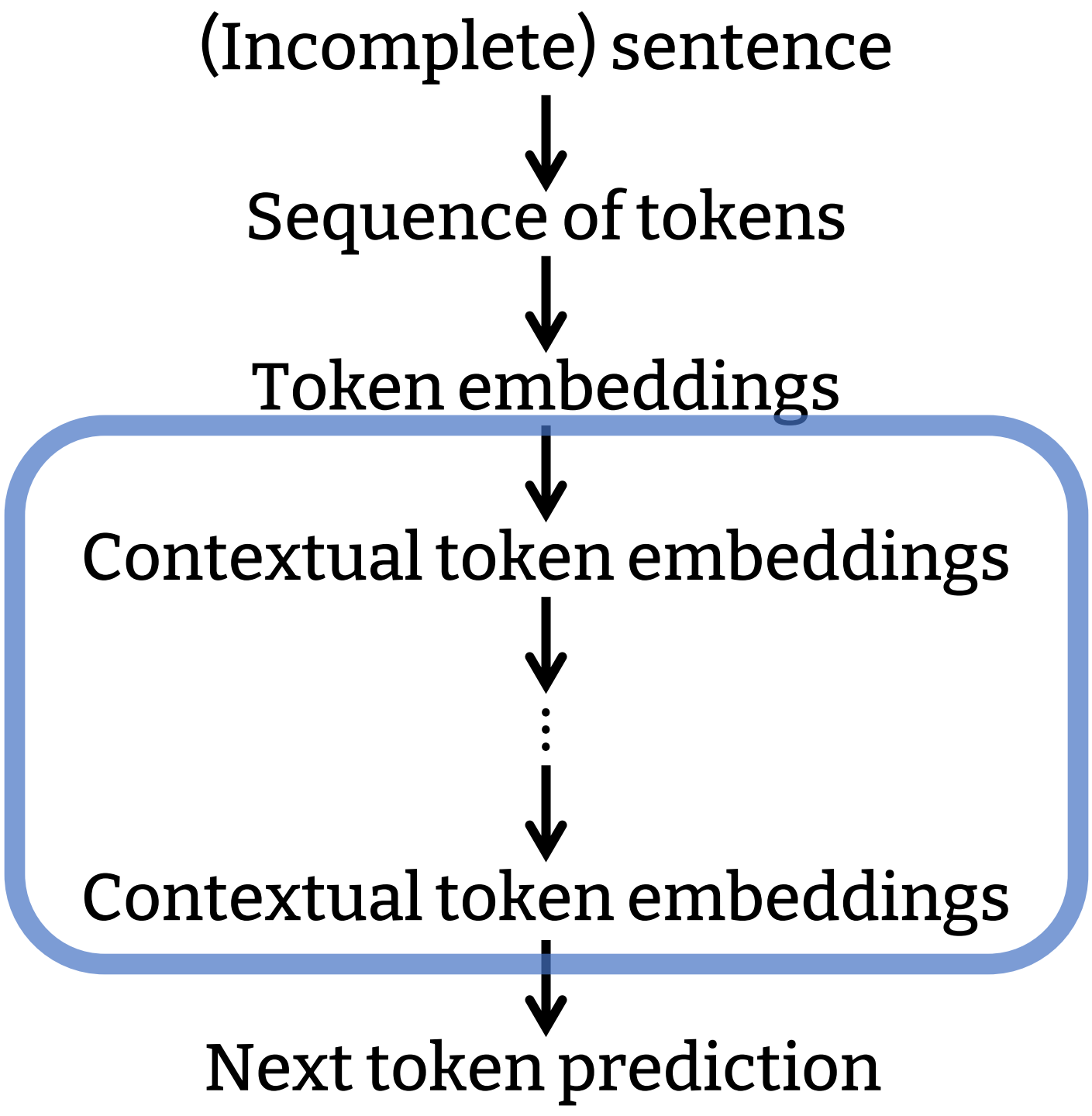

Alice was
a
the
in
not
born
also
an
still
so
just
Alice was an
American
early
amazing
old
excellent
avid
important
artist
extremely
accomplished
Alice was an important
part
figure
player
member
person
character
role
and
piece
source
in
,
.
to
and
for
on
who
of
that
the
The
this
my
a
many
both
our
Star
that
of
stories
films
movies
novels
books
episodes
ways
popular
other
the
our
these
my
those
her
his
Marvel
them
us
stories
films
original
movies
novels
early
books
first
games
series
of
,
.
that
I
in
and
we
she
about
've
read
have
told
wrote
know
tell
're
saw
had
seen
written
read
told
heard
in
.
,
today
now
in
.
,
so
of
from
and
on
here
since
the
this
The
our
comics
recent
Star
Marvel
movies
DC
series
comics
books
past
last
comic
show
book
Marvel
novels
few
couple
two
decade
year
three
20
five
several
10
years
decades
or
books
seasons
and
novels
months
to
-
.
,
,"
and
."
of
or
about
in
as
She
I
He
And
We
It
But
The
In
0.00
0.25
0.50
0.75
1.00


(Incomplete) sentence
Sequence of words
Word embeddings
Contextual word embeddings
Contextual word embeddings
Next word prediction


(Incomplete) sentence
Sequence of tokens
Token embeddings
Contextual token embeddings
Contextual token embeddings
Next token prediction



Progressively refine the embeddings



MLP
+
Attention
("transformers")

Text

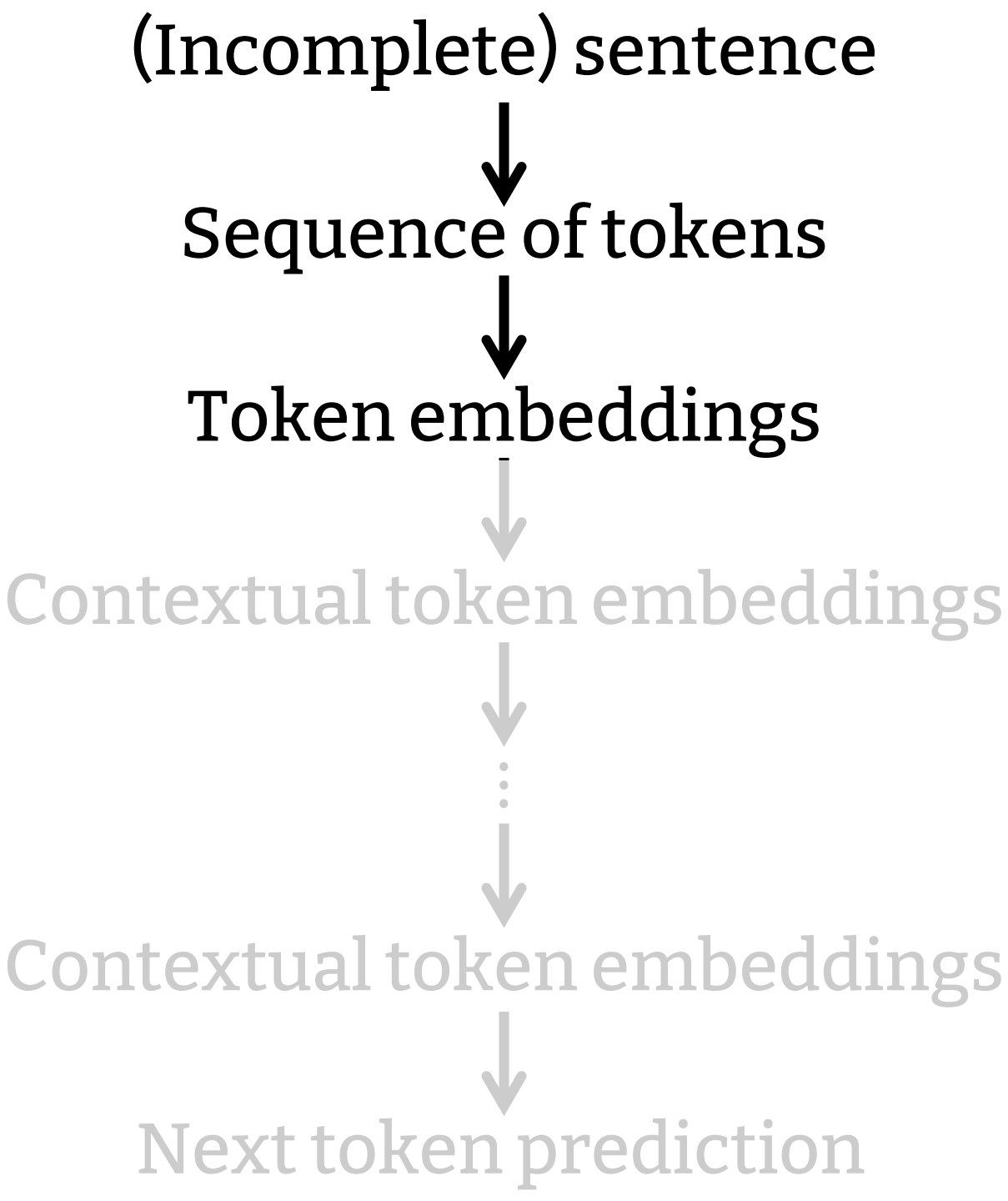
It was the best of times, it was the worst of times, it was the age of wisdom, it was the age of foolishness, it was the epoch of belief, it was the epoch of incredulity, it was the season of Light, it was the season of Darkness, it was the spring of hope, it was the winter of despair, we had everything before us, we had nothing before us, we were all going direct to Heaven, we were all going direct the other way--in short, the period was so far like the present period, that some of its noisiest authorities insisted on its being received, for good or for evil, in the superlative degree of comparison only.

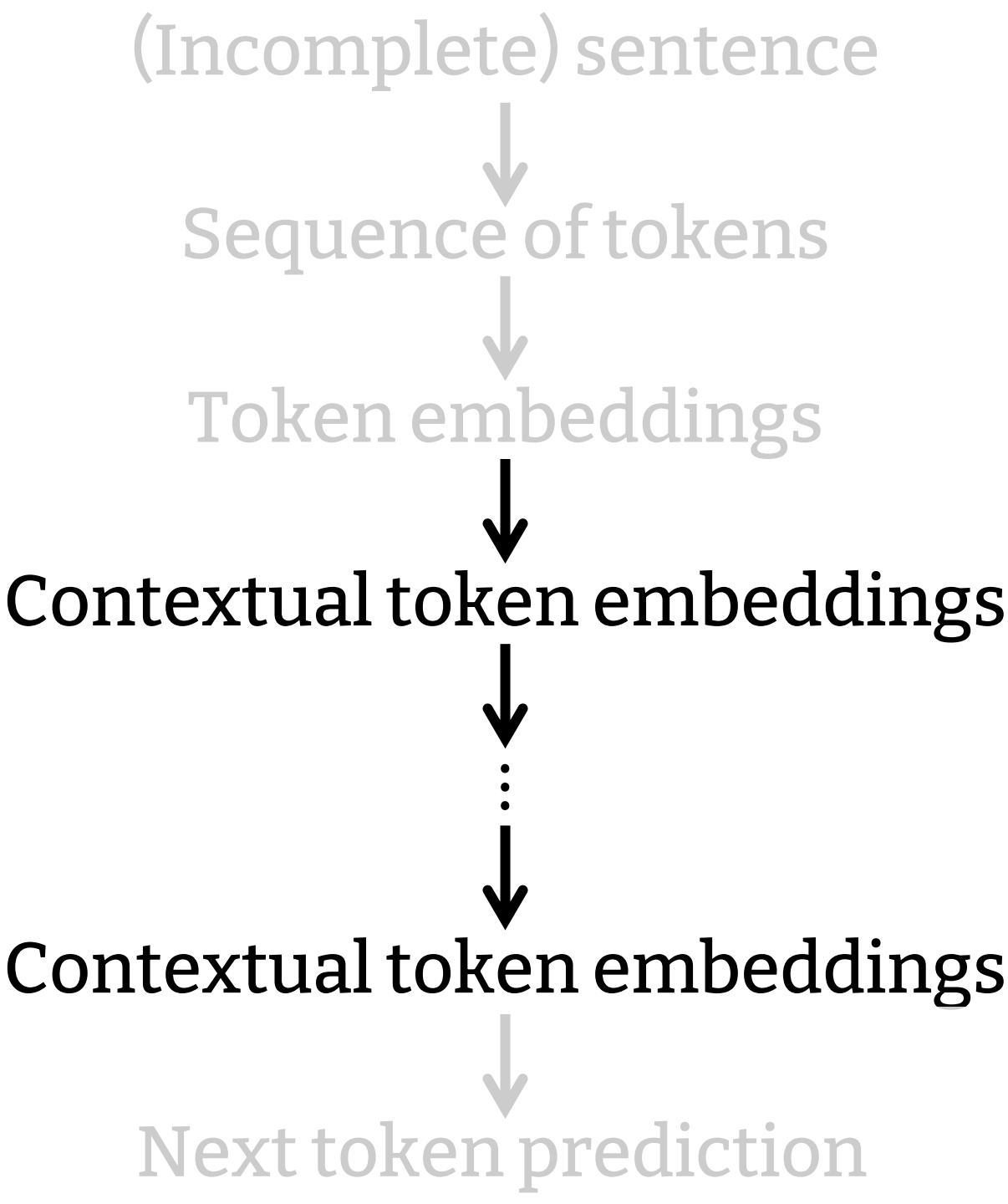
Tokens

It was the best of times, it was the worst of times, it was the age of wisdom, it was the age of foolishness, it was the epoch of belief, it was the epoch of incredulity, it was the season of Light, it was the season of Darkness, it was the spring of hope, it was the winter of despair, we had everything before us, we had nothing before us, we were all going direct to Heaven, we were all going direct the other way--in short, the period was so far like the present period, that some of its noisiest authorities insisted on its being received, for good or for evil, in the superlative degree of comparison only.

Token embeddings

```
"It"      [0.039, -0.087, 0.066, -0.053, -0.088, -0.069, -0.217, -0.041, 0.047, -0.003, …]
" was"    [-0.074, -0.084, 0.181, -0.107, -0.085, 0.056, -0.282, -0.035, -0.104, 0.073, …]
" the"    [-0.039, 0.005, 0.042, 0.04, -0.036, -0.055, -0.257, -0.002, 0.025, 0.043, …]
" best"   [-0.151, 0.055, 0.026, 0.071, 0.012, 0.041, -0.319, 0.011, 0.059, -0.269, …]
" of"     [-0.057, 0.018, 0.033, 0.041, 0.012, -0.04, -0.253, 0.002, 0.065, 0.066, …]
" times"  [-0.037, -0.02, 0.107, -0.167, -0.011, 0.094, -0.284, 0.183, 0.082, -0.028, …]
","       [0.011, -0.003, 0.032, 0.055, 0.052, -0.06, -0.24, -0.017, 0.039, 0.04, …]
" it"     [0.026, -0.047, 0.033, 0.051, -0.058, -0.026, -0.259, 0.01, 0.043, 0.047, …]
" was"    [-0.074, -0.084, 0.181, -0.107, -0.085, 0.056, -0.282, -0.035, -0.104, 0.073, …]
" the"    [-0.039, 0.005, 0.042, 0.04, -0.036, -0.055, -0.257, -0.002, 0.025, 0.043, …]
" worst"  [-0.051, -0.0, 0.253, 0.063, -0.003, -0.043, -0.343, -0.124, -0.111, -0.048, …]
" of"     [-0.057, 0.018, 0.033, 0.041, 0.012, -0.04, -0.253, 0.002, 0.065, 0.066, …]
" times"  [-0.037, -0.02, 0.107, -0.167, -0.011, 0.094, -0.284, 0.183, 0.082, -0.028, …]
","       [0.011, -0.003, 0.032, 0.055, 0.052, -0.06, -0.24, -0.017, 0.039, 0.04, …]
```


(Incomplete) sentence
Sequence of tokens
Token embeddings
Contextual token embeddings
Contextual token embeddings
Next token prediction


(Incomplete) sentence
Sequence of tokens
Token embeddings
Contextual token embeddings
Contextual token embeddings
Next token prediction

Token embedding

bank

↓

```
[0.046, 0.082, 0.137, 0.038, -0.027, 0.06, -0.344, -0.193, 0.08, 0.226, …]
```

Token embedding

bank

↓

Token embeddings

central bank river bank power bank

Contextual embeddings

central bank

river bank

power bank









(Incomplete) sentence

↓

Sequence of tokens

↓

Token embeddings

↓

Contextual token embeddings

↓

⋮

↓

Contextual token embeddings

↓

Next token prediction

MLP

+

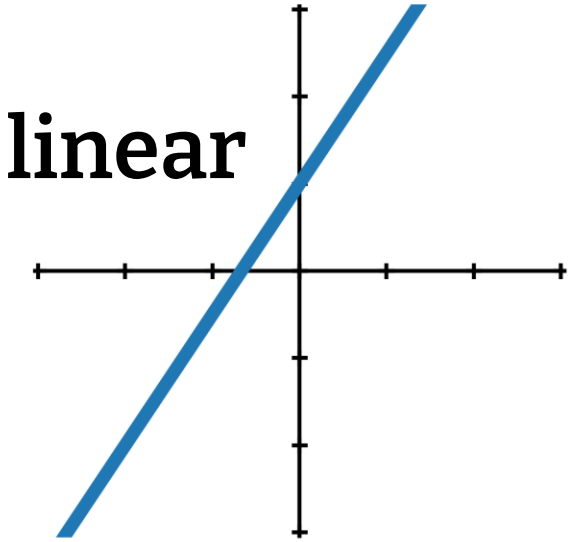
Attention

(“transformers”)

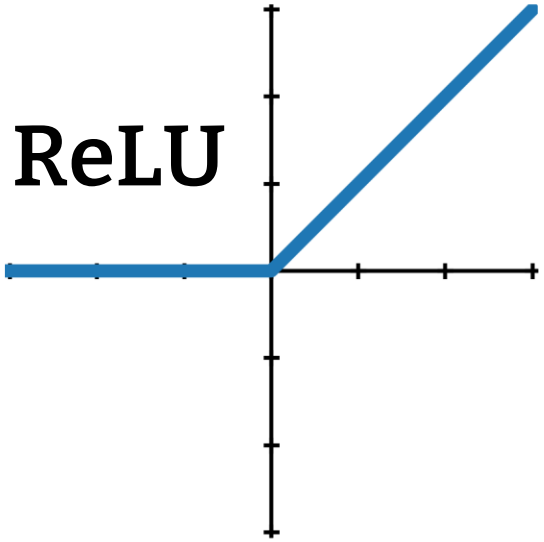
MLP: one layer, example

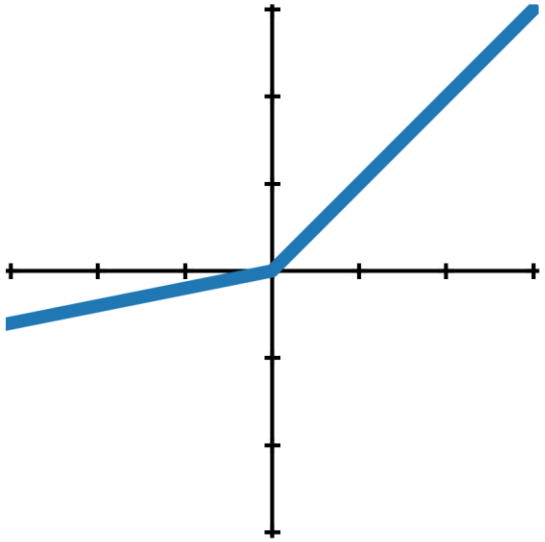
$$x \longmapsto y = 3x_1+7x_2-x_3+2x_4+\cdots \longmapsto z = \sigma(y)$$

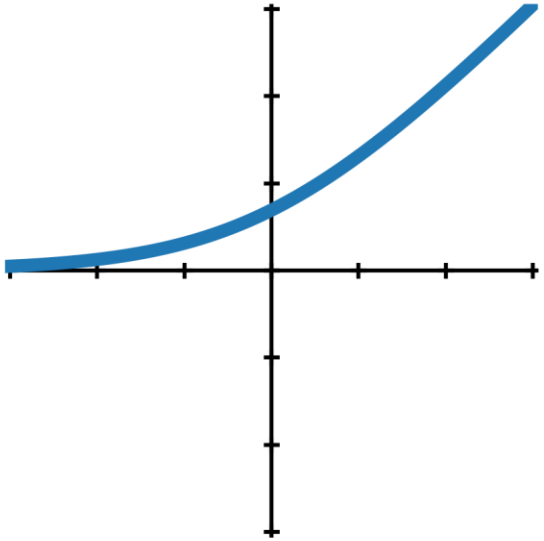
Linear functions

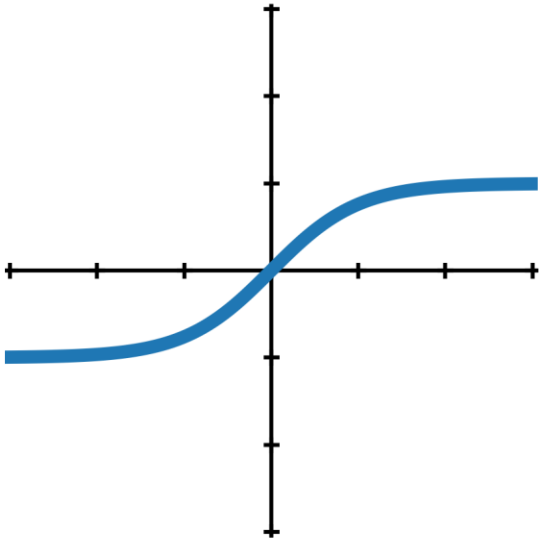


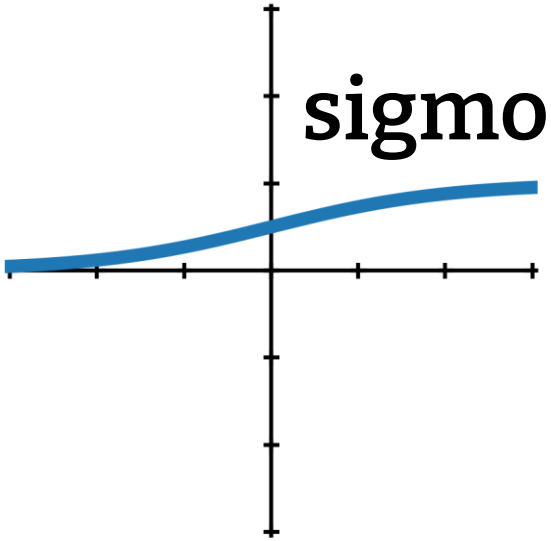
Non-linear functions σ











MLP: one layer, example

$$x \longmapsto y = 3x_1+7x_2-x_3+2x_4+\cdots \longmapsto z = \sigma(y)$$

MLP: one layer

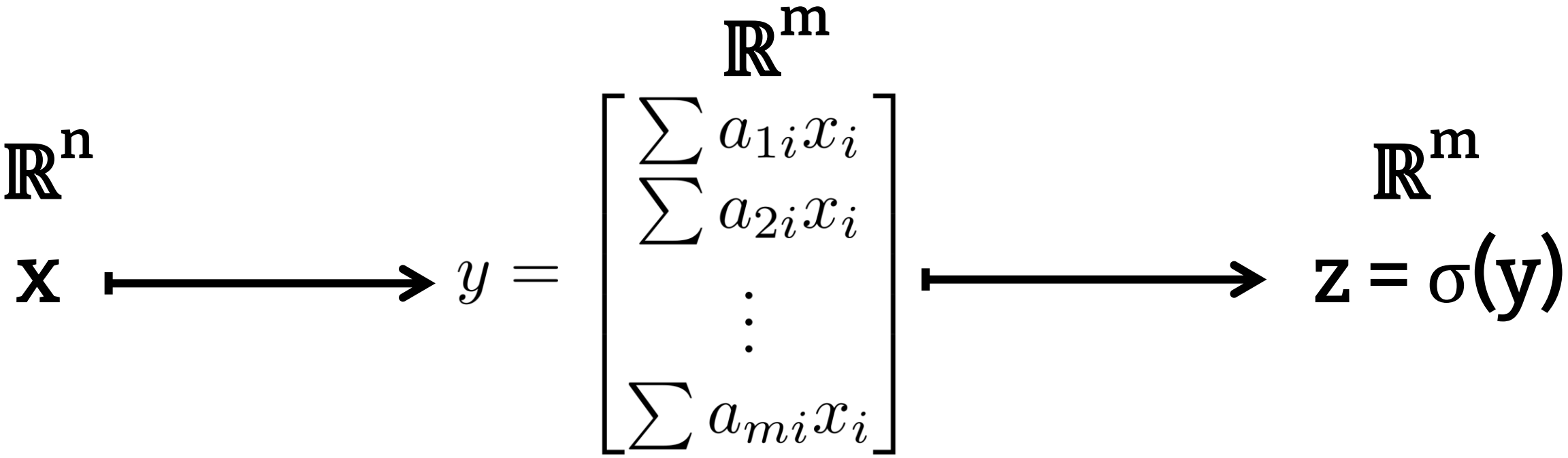
$$x \longmapsto y = \sum a_i x_i \longmapsto z = \sigma(y)$$

MLP: one layer

$$\mathbb{R}^n \qquad \mathbb{R} \qquad \mathbb{R}$$

$$x \longmapsto y = \sum a_i x_i \longmapsto z = \sigma(y)$$

MLP: one layer

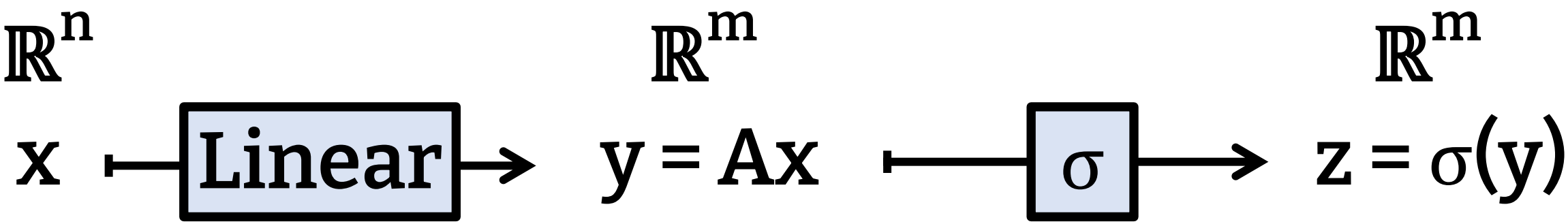


MLP: one layer

$$\mathbb{R}^n \qquad\qquad \mathbb{R}^m \qquad\qquad \mathbb{R}^m$$

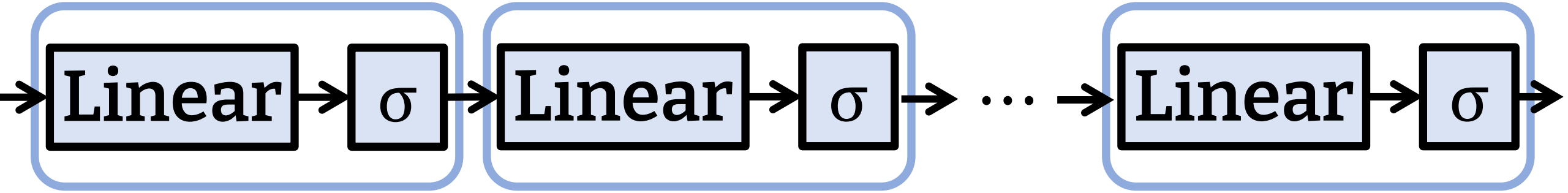
$$x \longmapsto y = Ax \longmapsto z = \sigma(y)$$

MLP: one layer

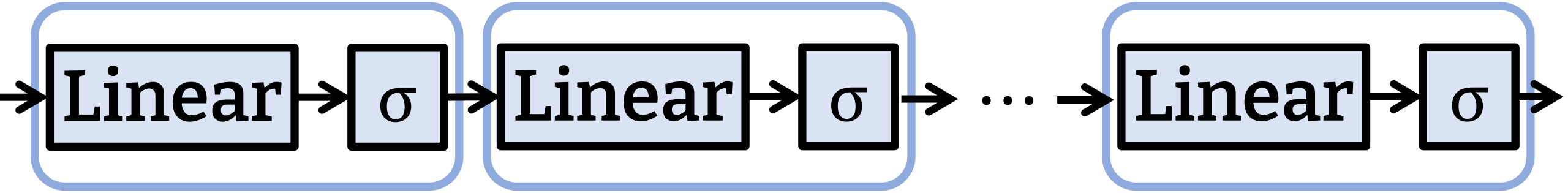


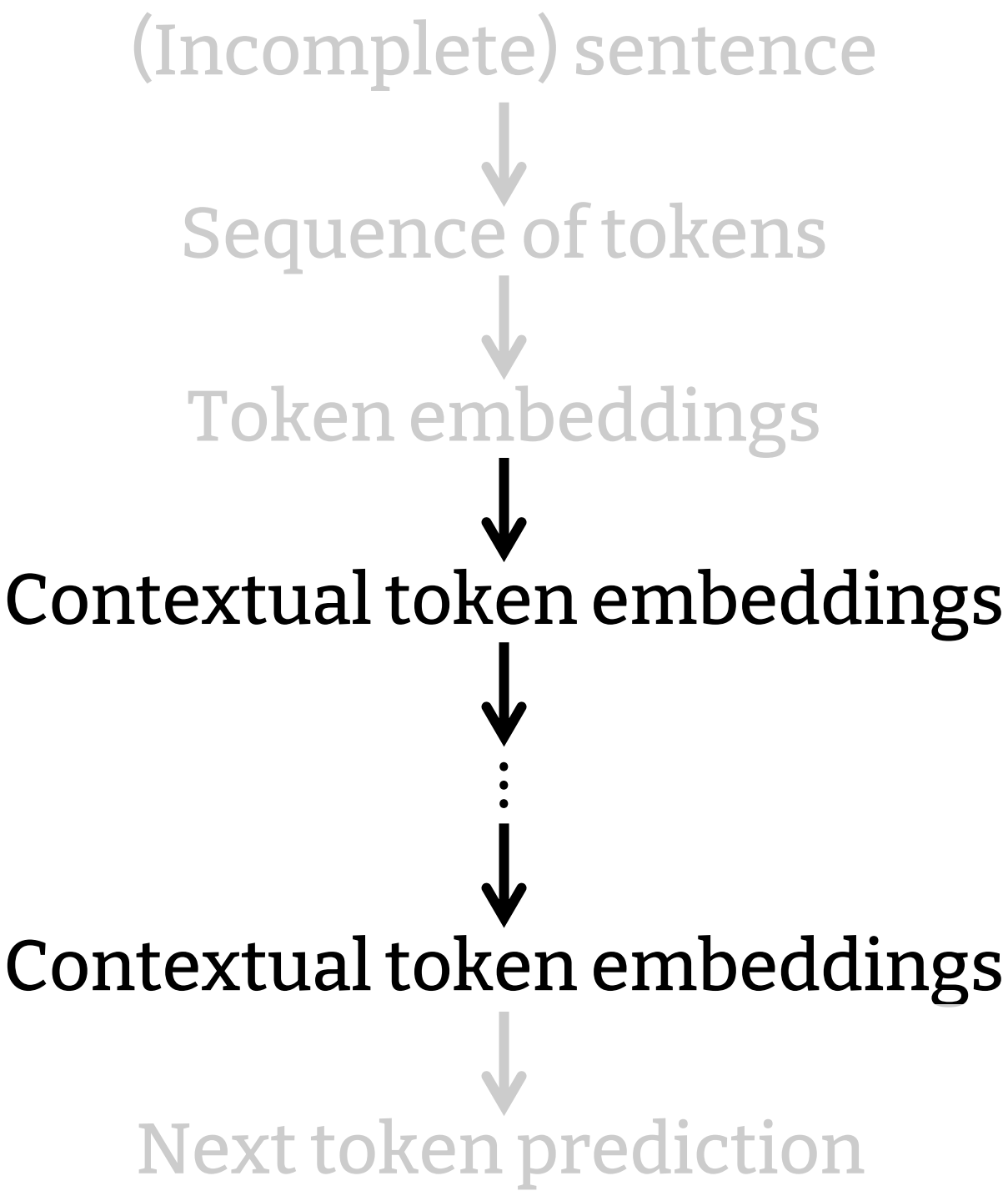
MLP: one layer

MLP: several layers



MLP (Multilayer Perceptron)





MLP

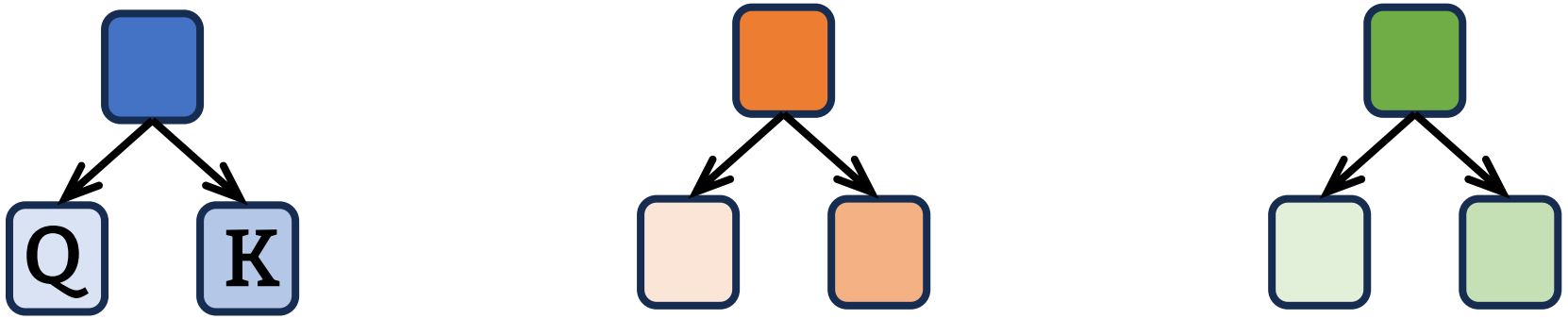
+

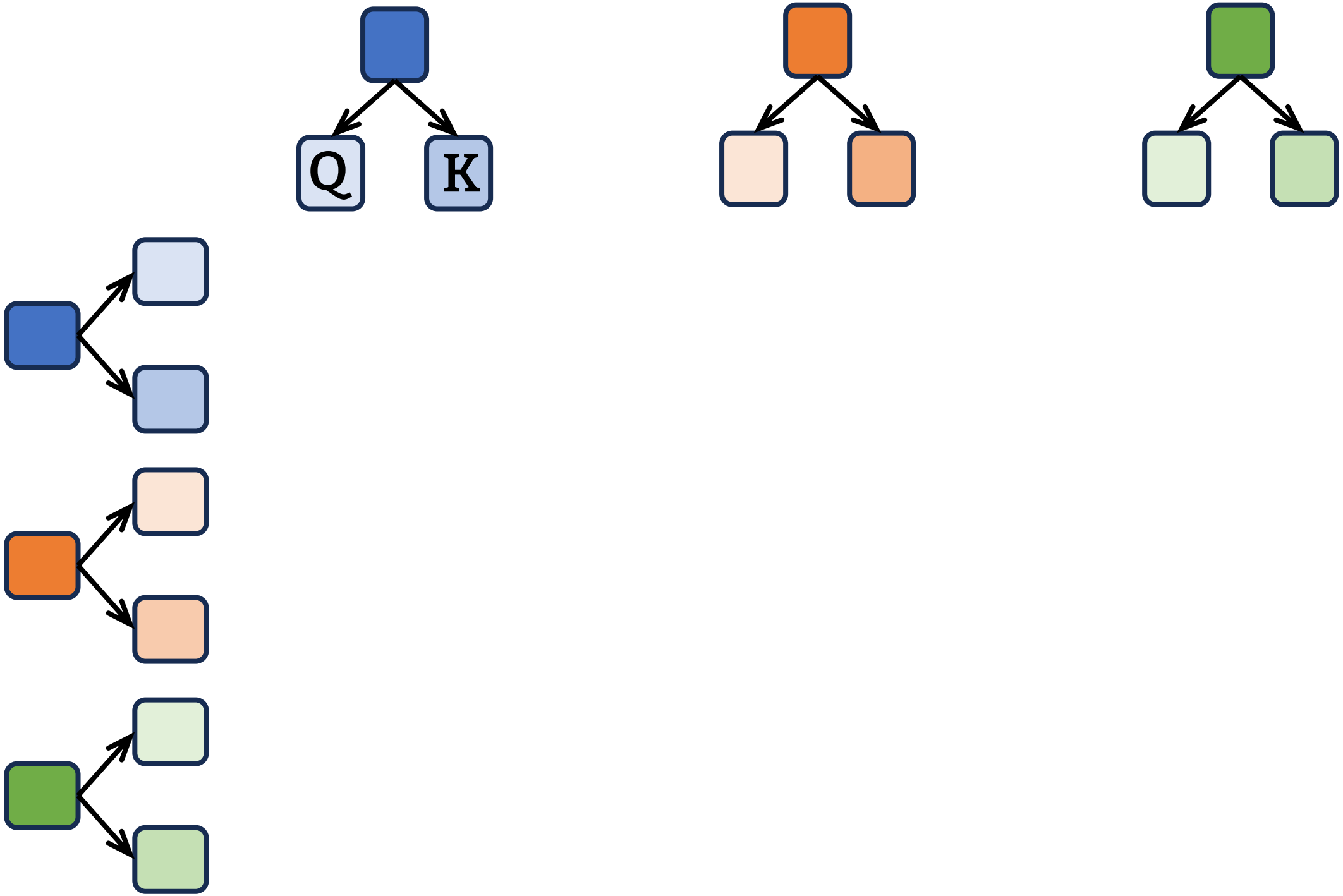
Attention

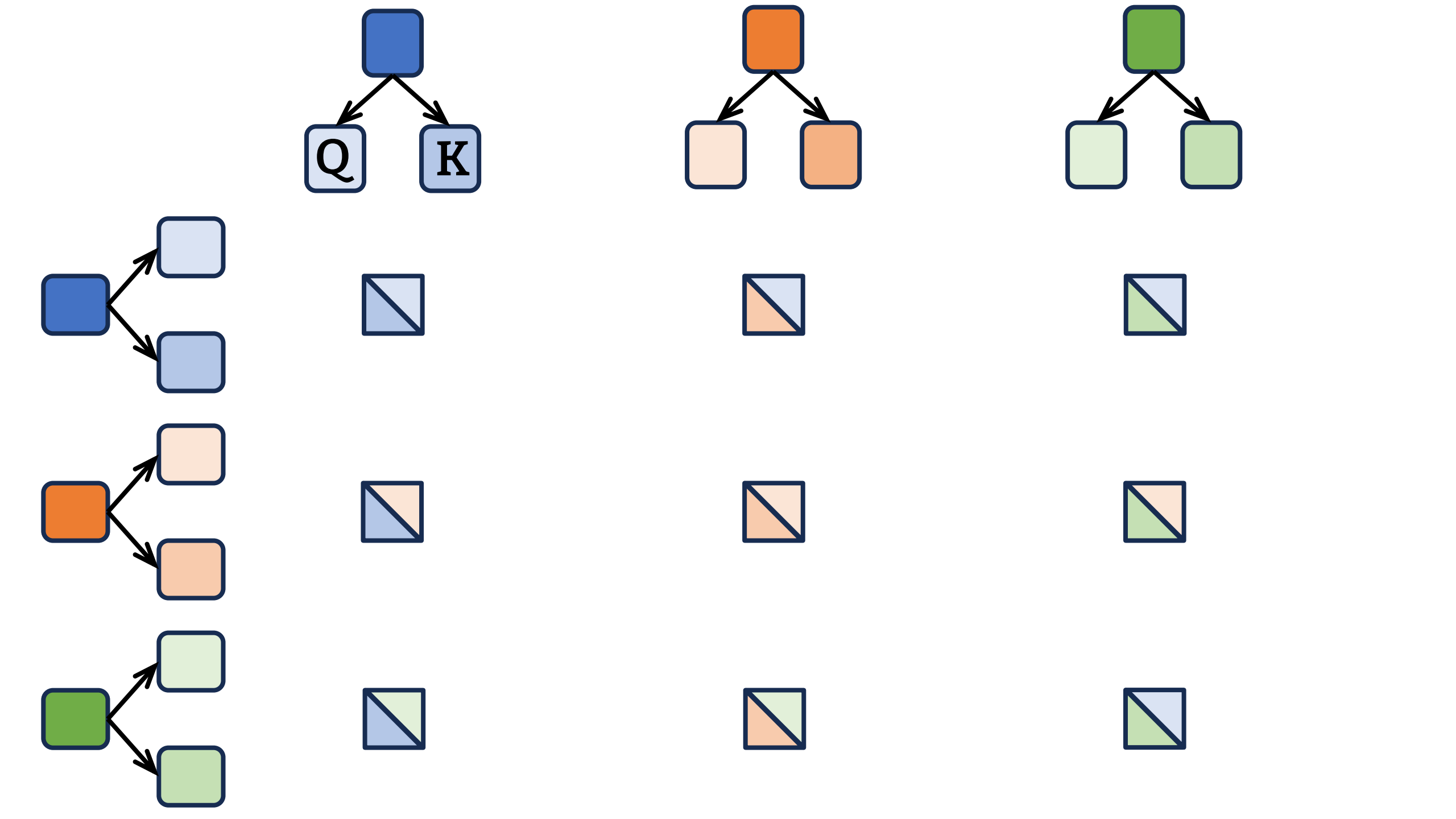
(“transformers”)

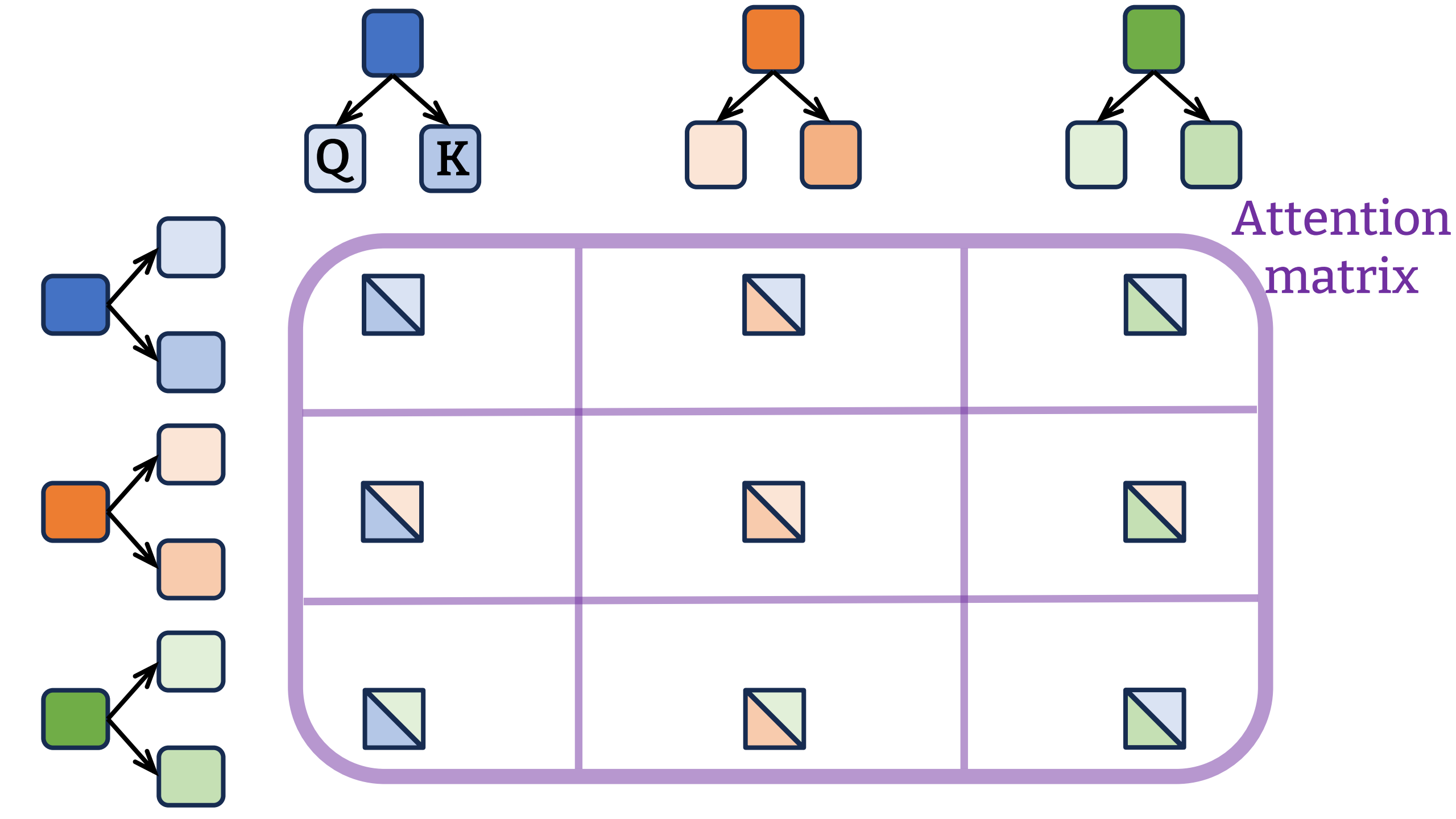
Attention mechanism

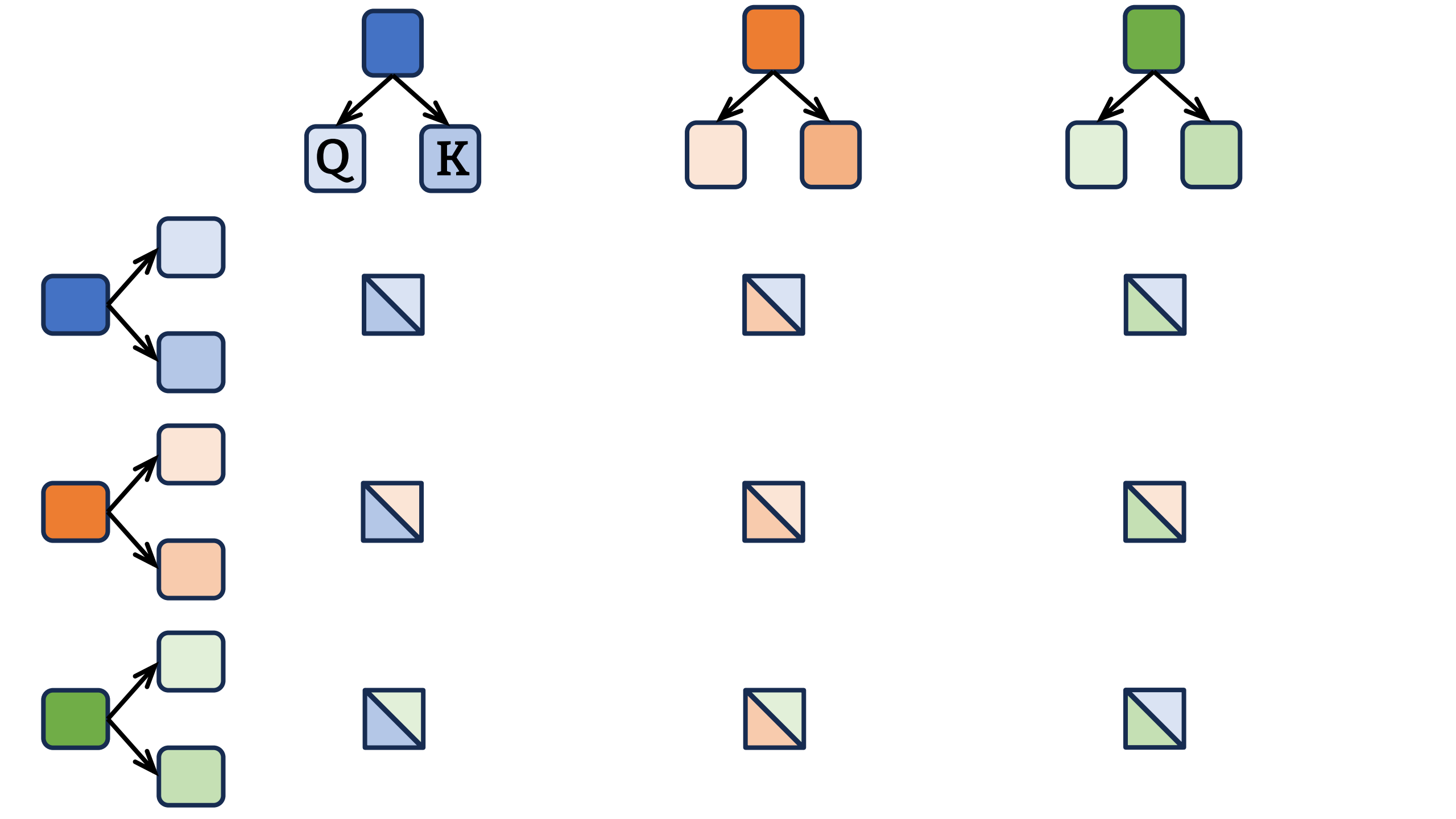
- MLP: weighted sums
- Attention: products

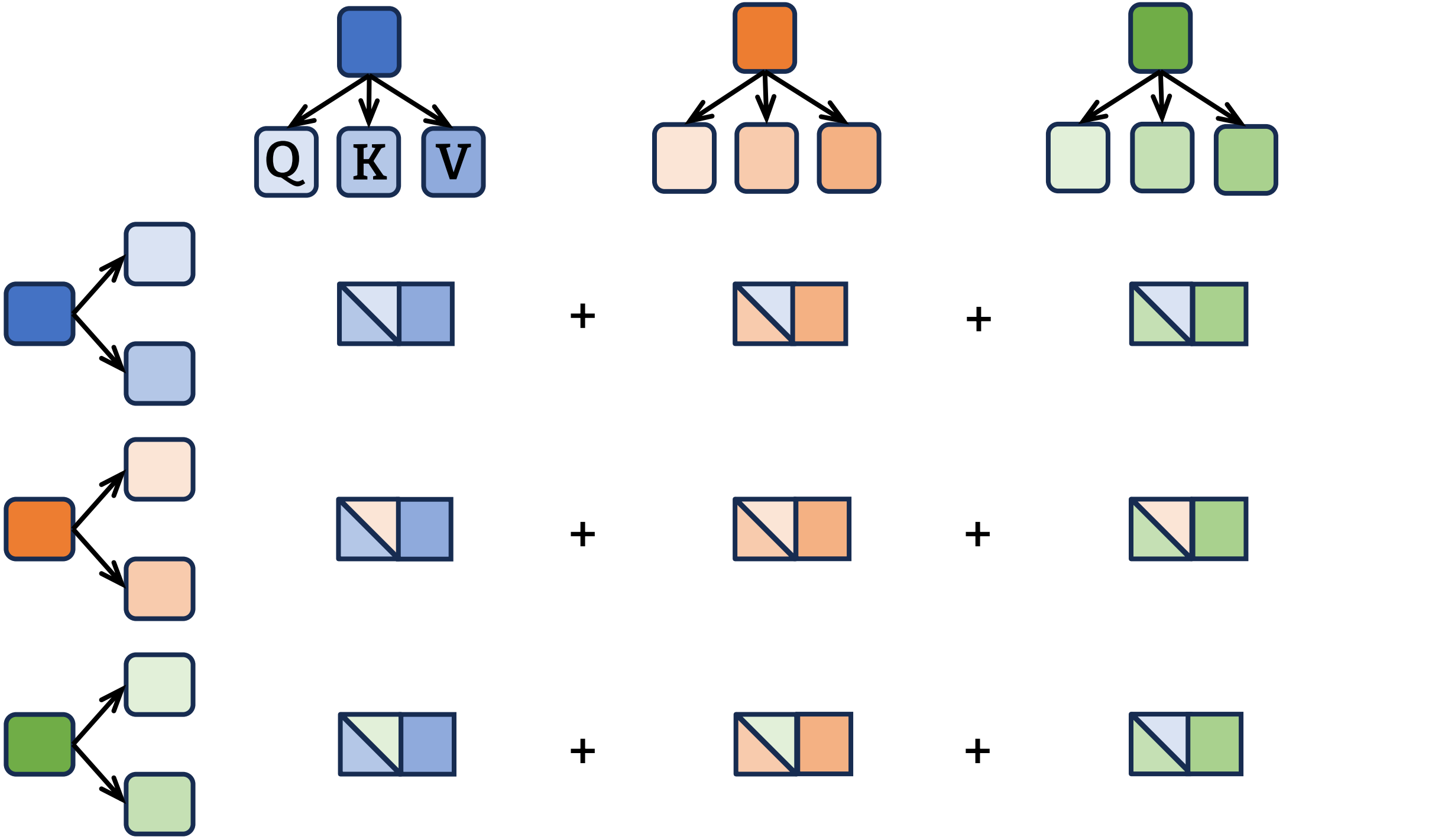

Q
K

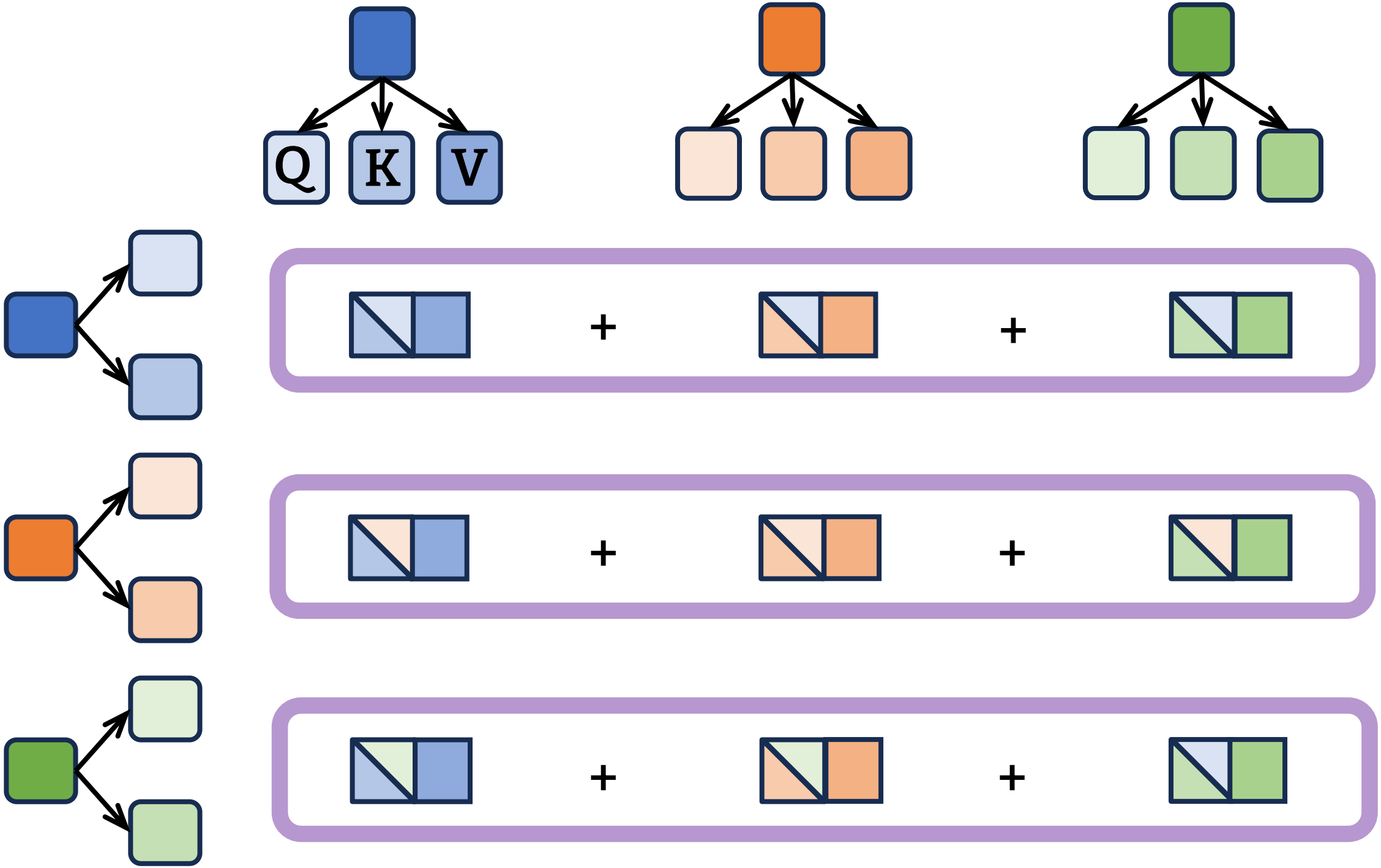

Q
K


Q
K

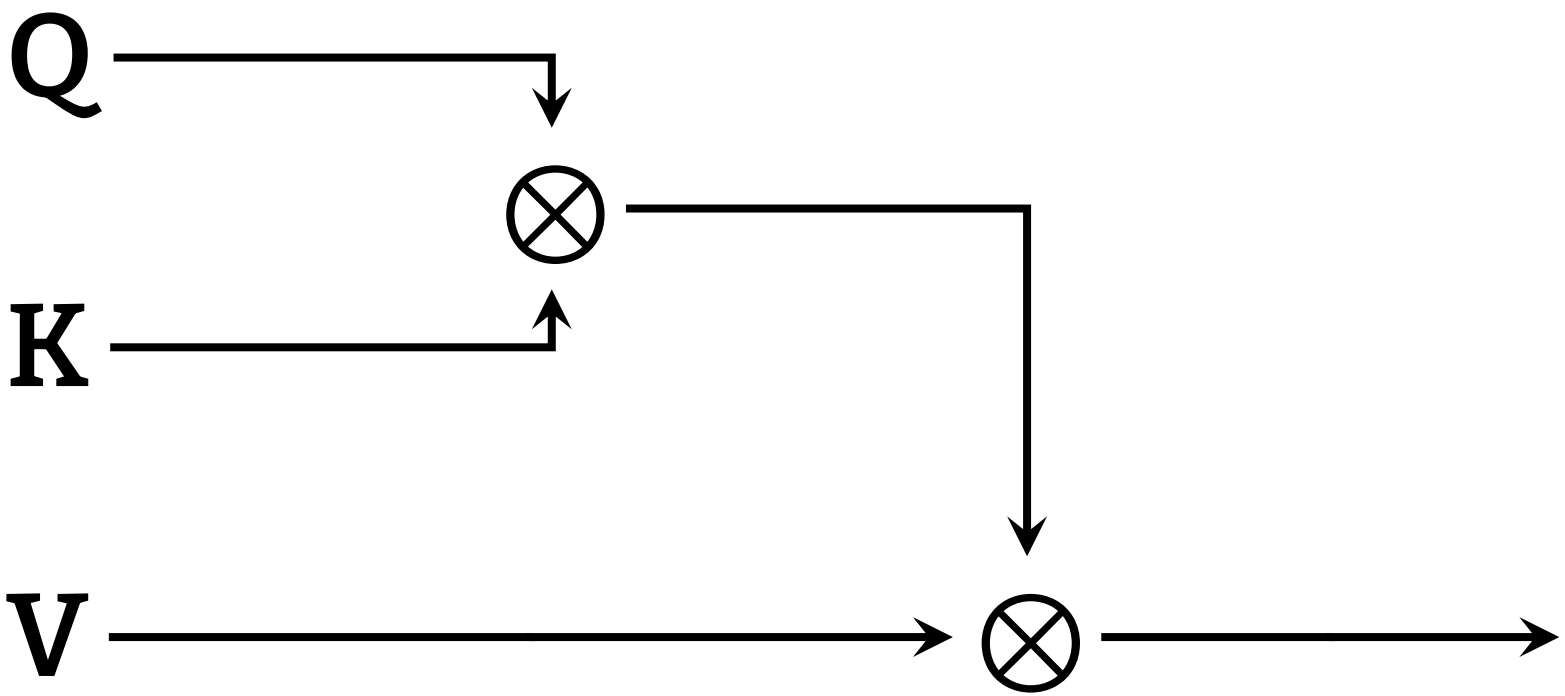

Q
K
Attention
matrix


Q
K

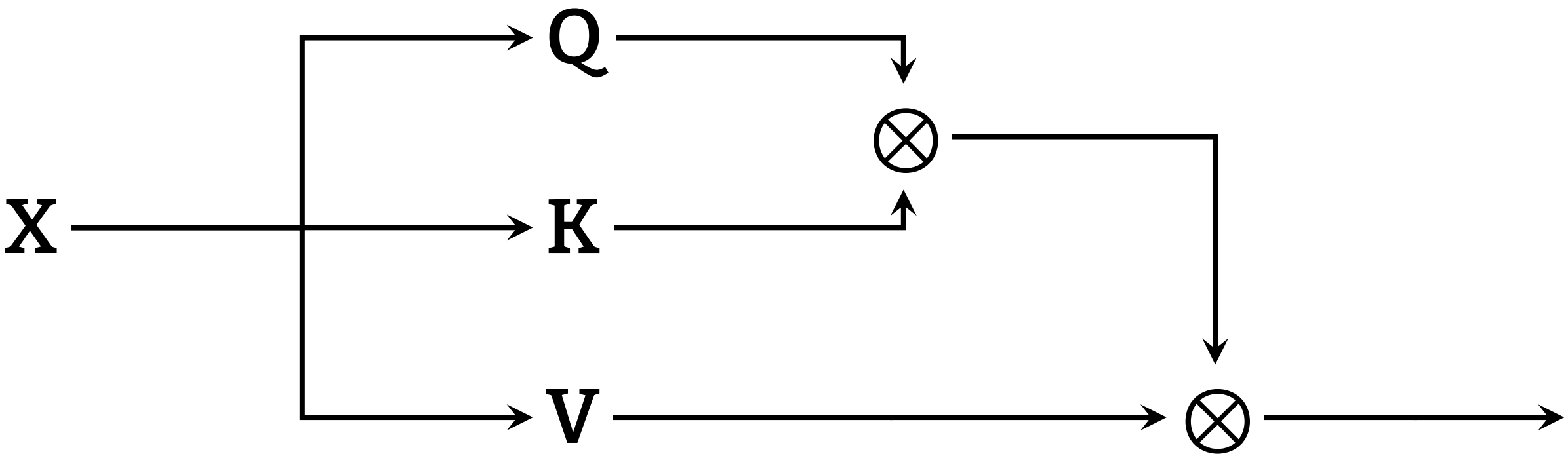

Q
K
V
+
+
+
+
+
+


Q
K
V
+
+
+
+
+
+

Attention

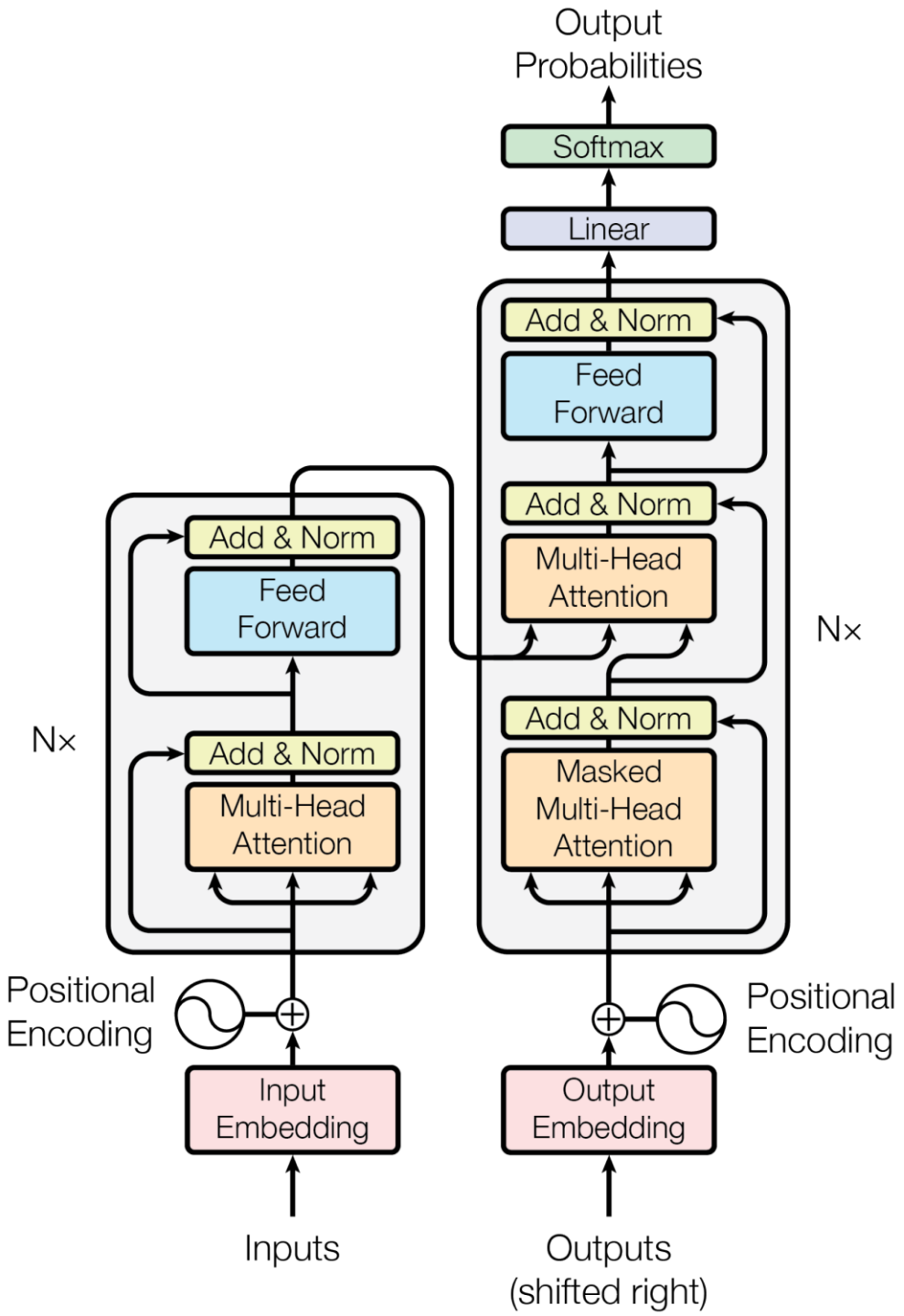


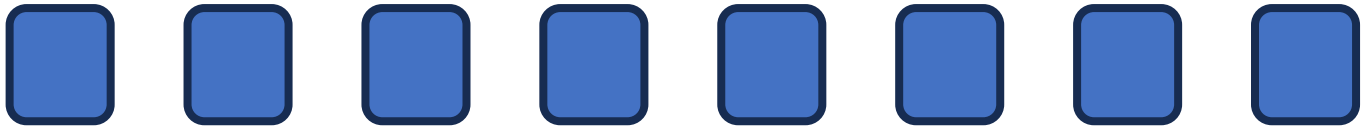
Self-Attention



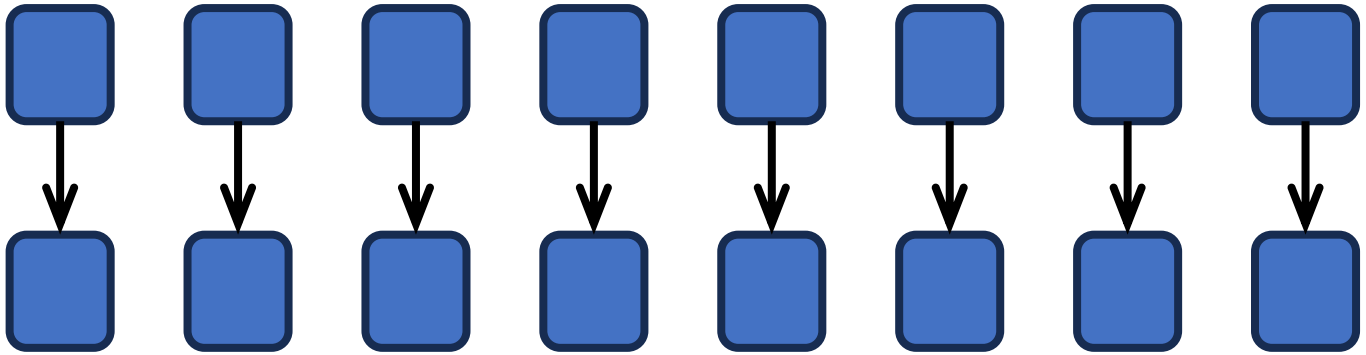
Transformer

- Transformer = attention + lots of details

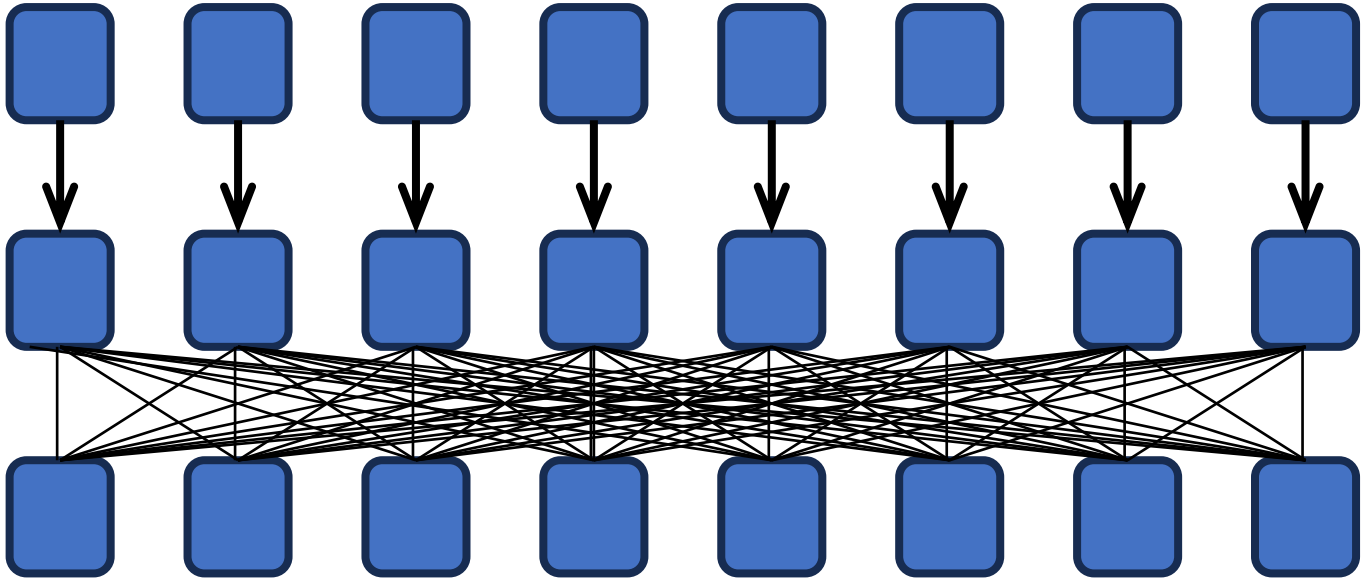




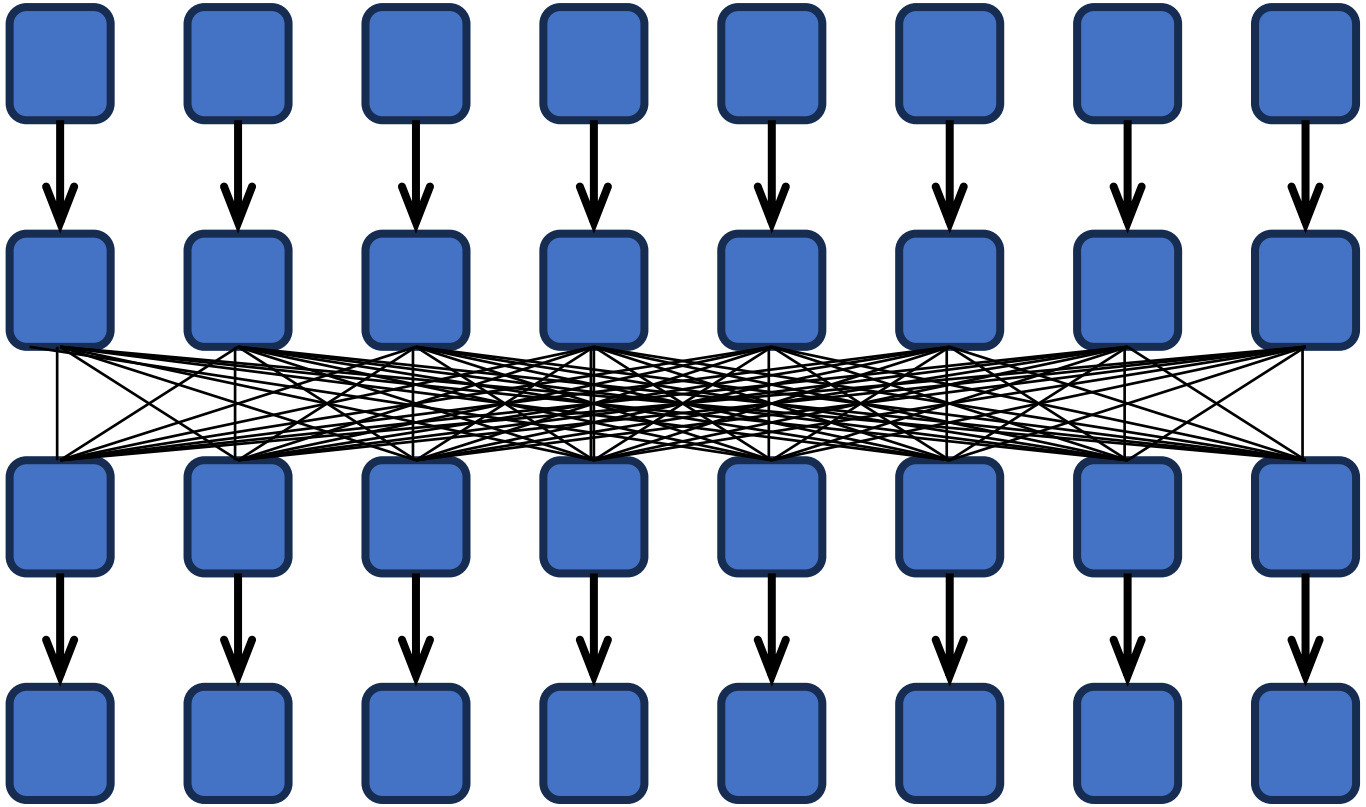
Tokens



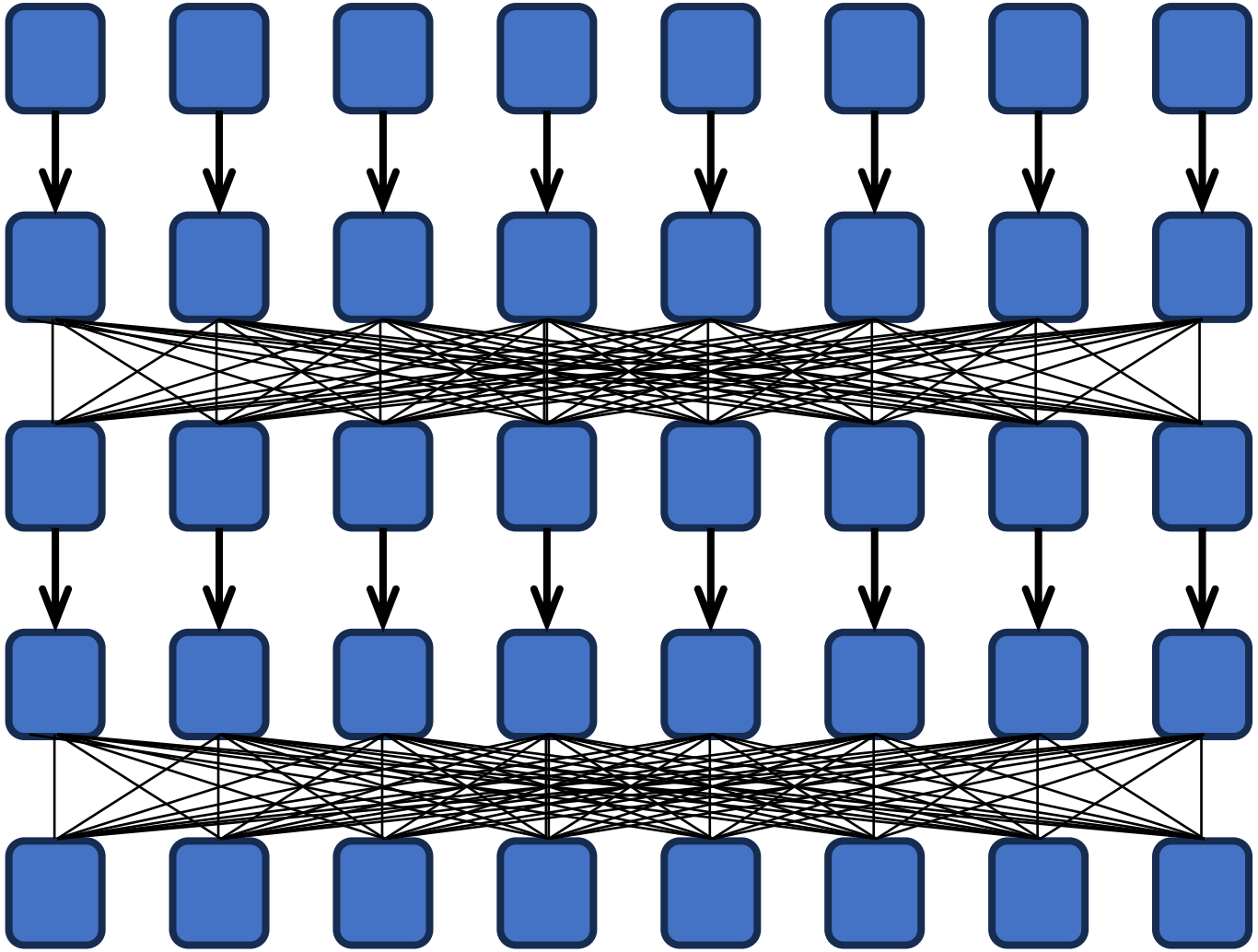
MLP



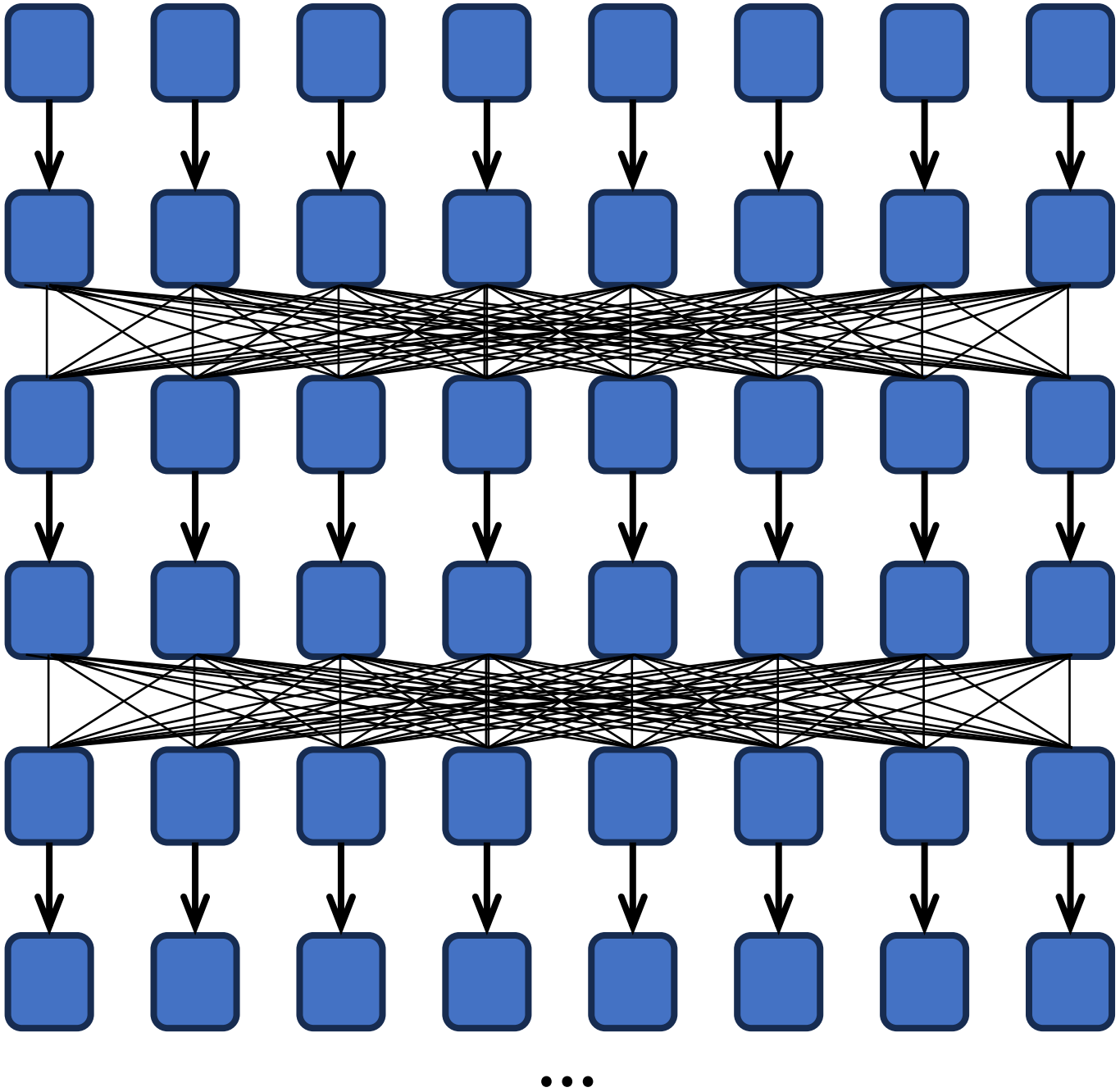
Attention

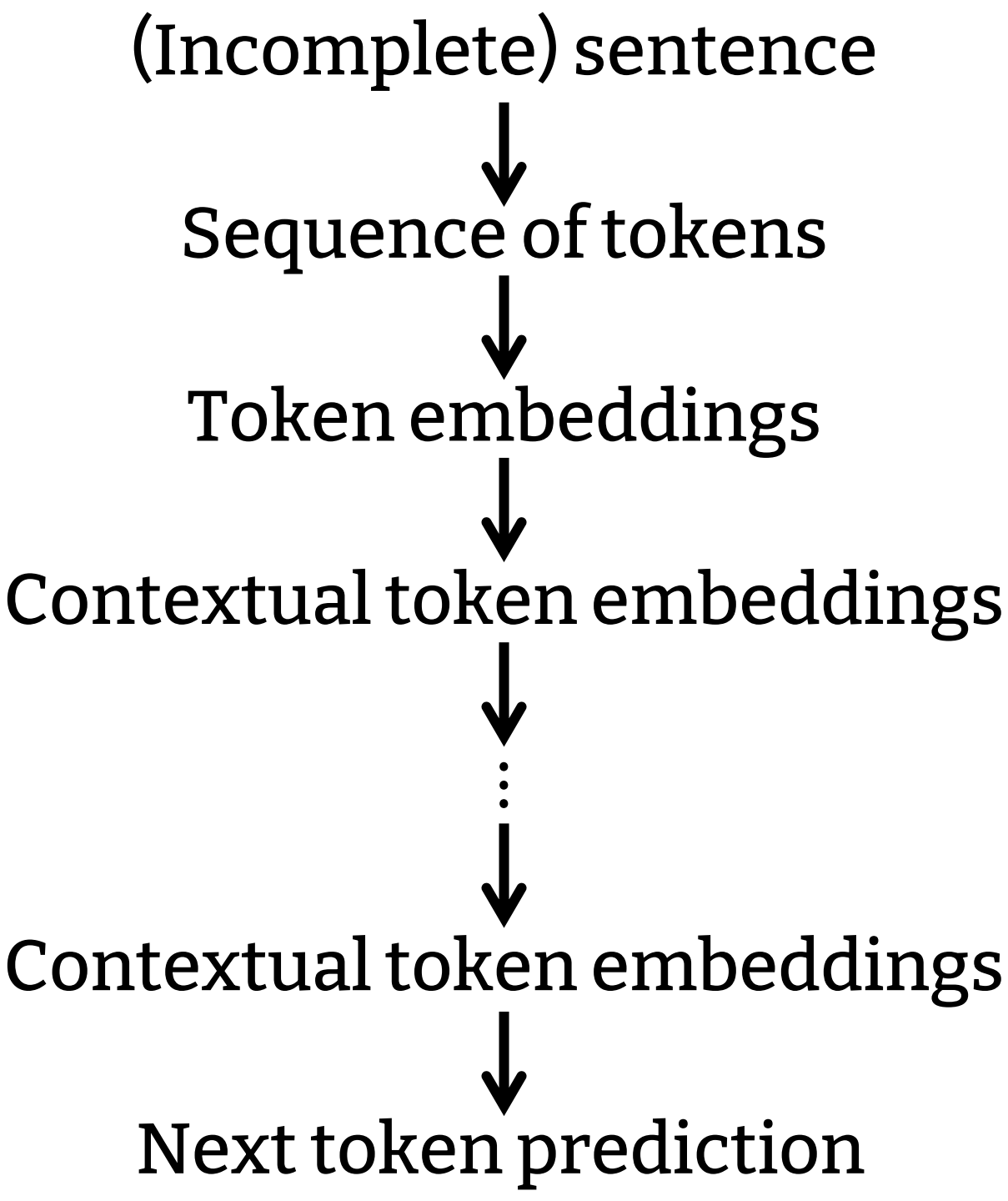


MLP



Attention


MLP
Attention
MLP
Attention
MLP
...


(Incomplete) sentence
Sequence of tokens
Token embeddings
Contextual token embeddings
Contextual token embeddings
Next token prediction

Not covered

- Training
- Fine-tuning: text completion → assistant
- Prompt engineering (few-shot learning, CoT)
- Tools (MCP)
- Agents

References

- 3blue1brown (YouTube):
  - Large language models explained briefly
  - Visualizing transformers and attention
  - Neural networks
- Andrej Karpathy (YouTube):
  - Deep dive into LLMs like ChatGPT

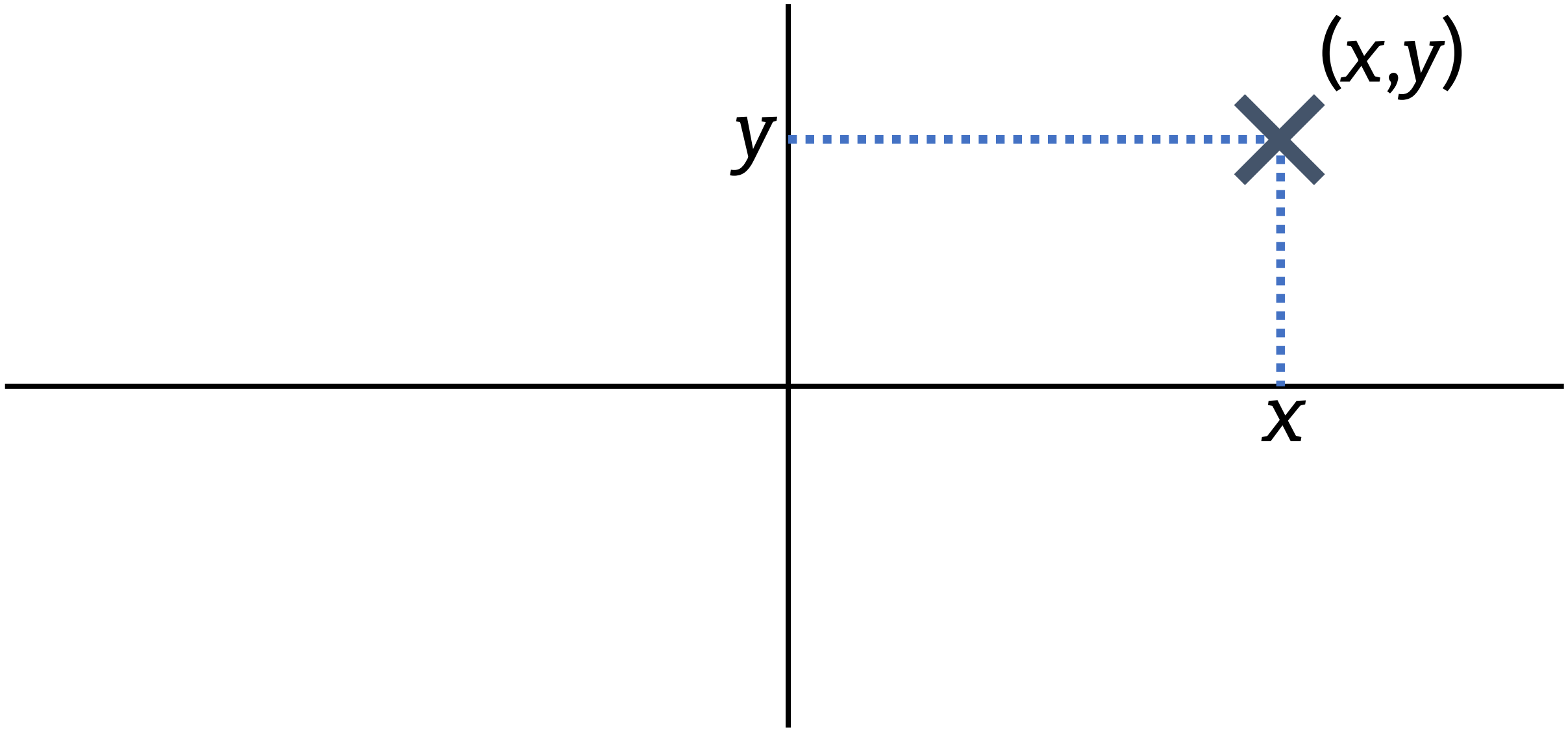
Extra Slides

Prerequisites

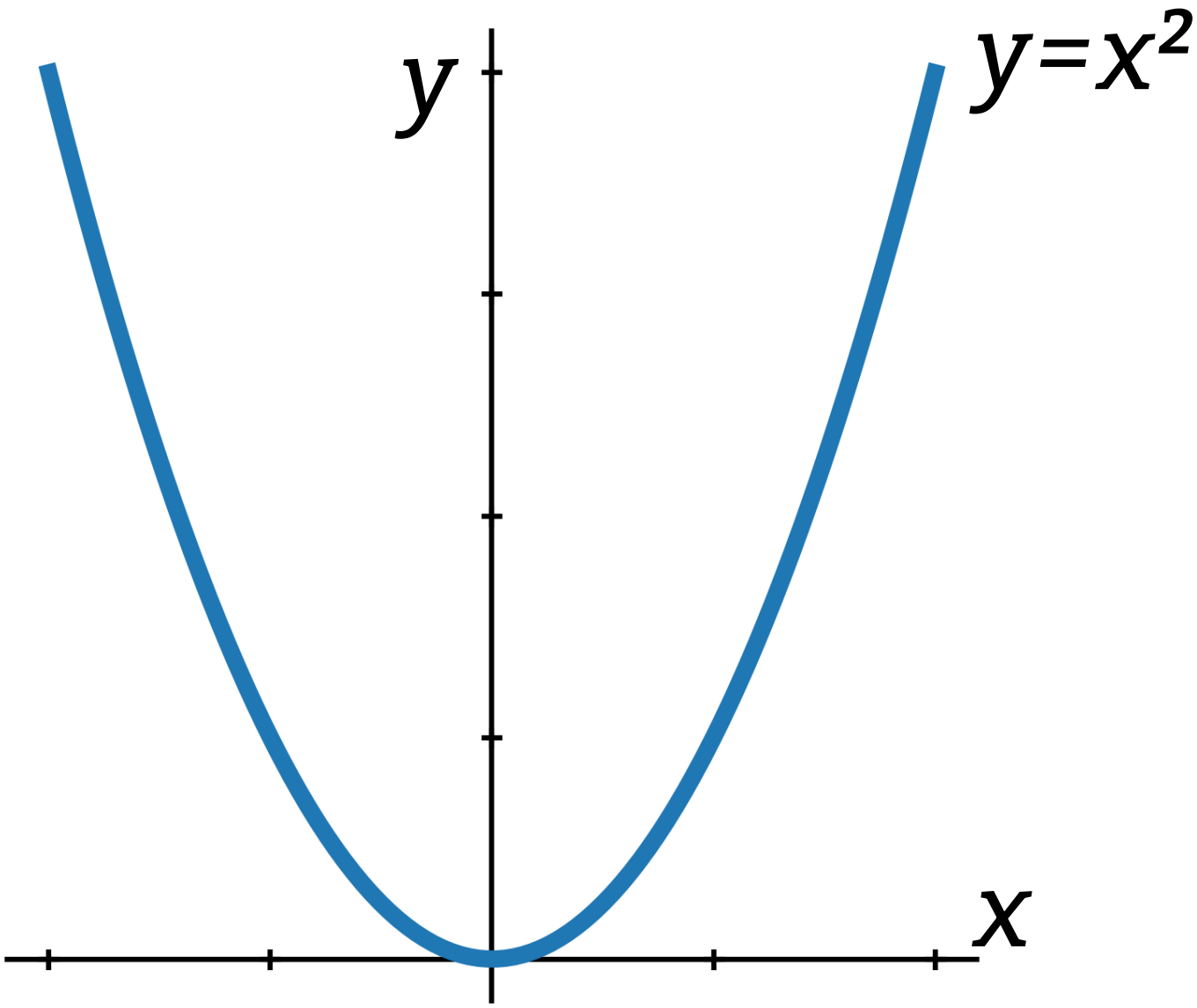
Function

$$f : \begin{cases} \mathbb{R} \longrightarrow \mathbb{R} \\ x \longmapsto x^2 \end{cases}$$

Coordinates



Graph of a function



Summation notation

$$\sum_{i=1}^{n} a_i = a_1 + a_2 + \cdots + a_n$$

Vectors

$$(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$$

Vectors

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbf{R}^n$$

Scalar product

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

Scalar product

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \sum_{i=1}^{n} x_i y_i$$

Scalar product

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \sum_i x_i y_i$$

Matrix

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix} \in \mathbf{R}^{n \times m}$$

Matrix product

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + \cdots + a_{1m}x_m \\ a_{21}x_1 + \cdots + a_{2m}x_m \\ \vdots \\ a_{n1}x_1 + \cdots + a_{nm}x_m \end{bmatrix}$$

Matrix product

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + \cdots + a_{1m}x_m \\ a_{21}x_1 + \cdots + a_{2m}x_m \\ \vdots \\ a_{n1}x_1 + \cdots + a_{nm}x_m \end{bmatrix}$$

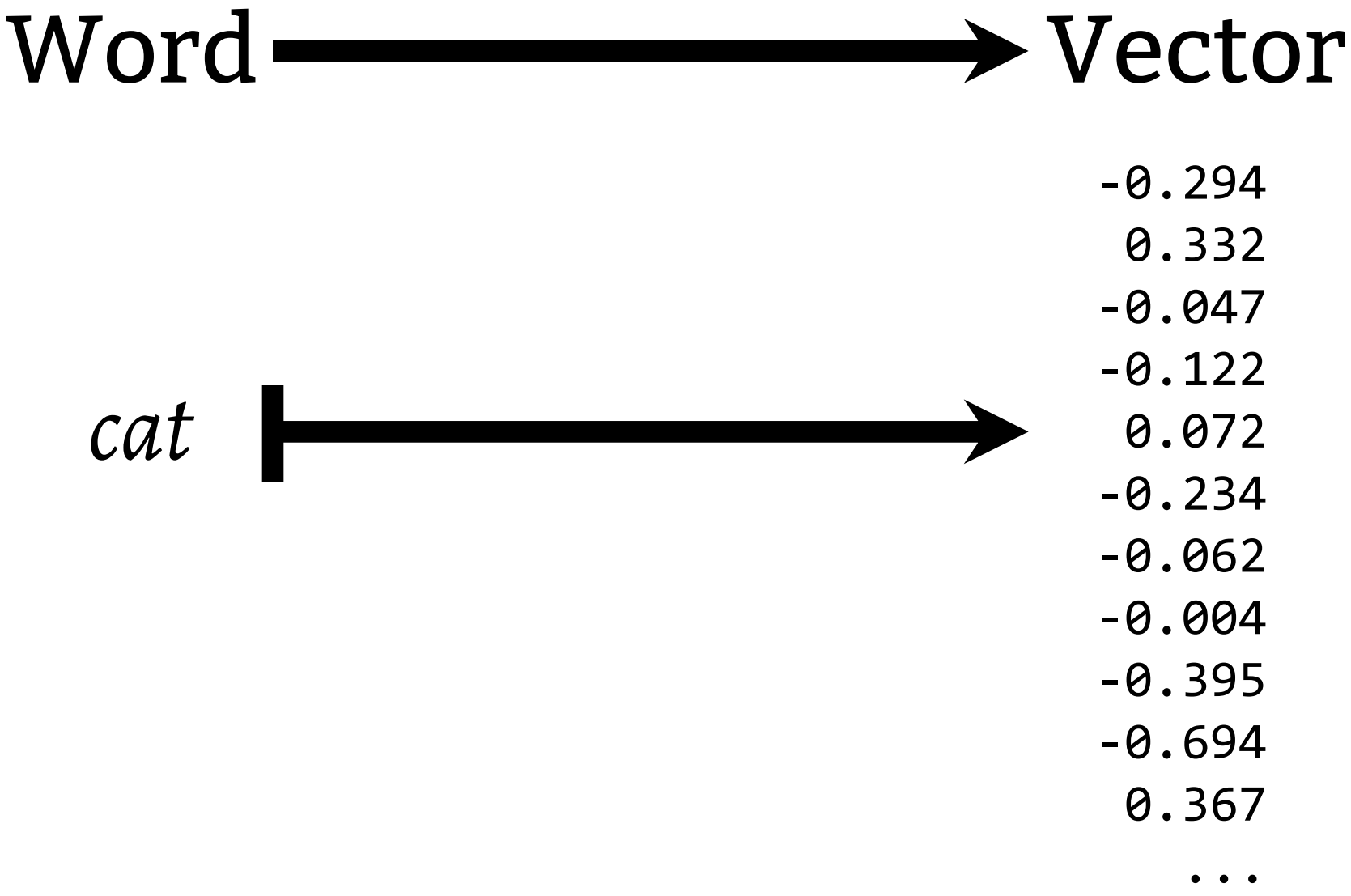
Matrix product

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + \cdots + a_{1m}x_m \\ a_{21}x_1 + \cdots + a_{2m}x_m \\ \vdots \\ a_{n1}x_1 + \cdots + a_{nm}x_m \end{bmatrix}$$

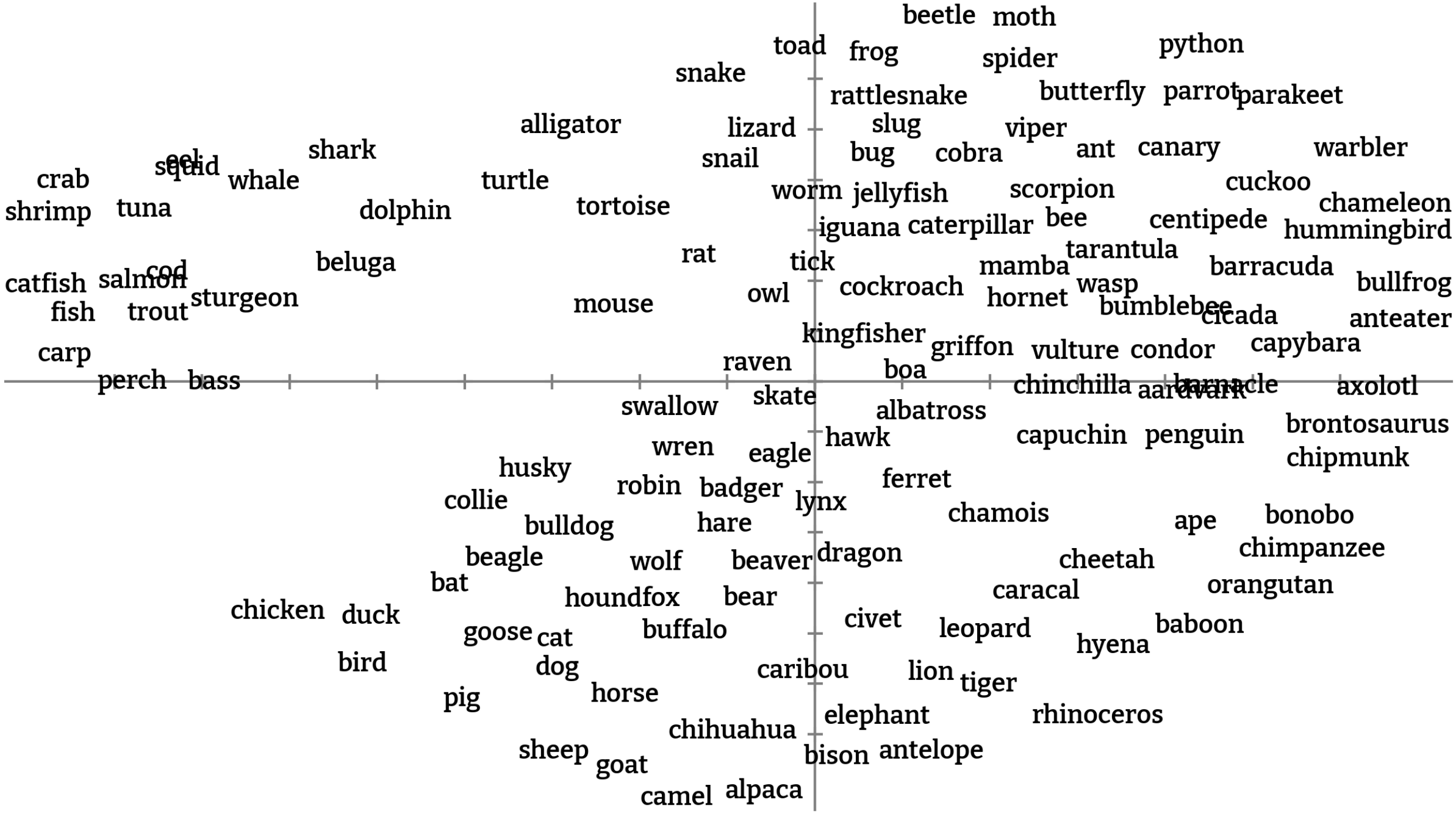
Matrix product

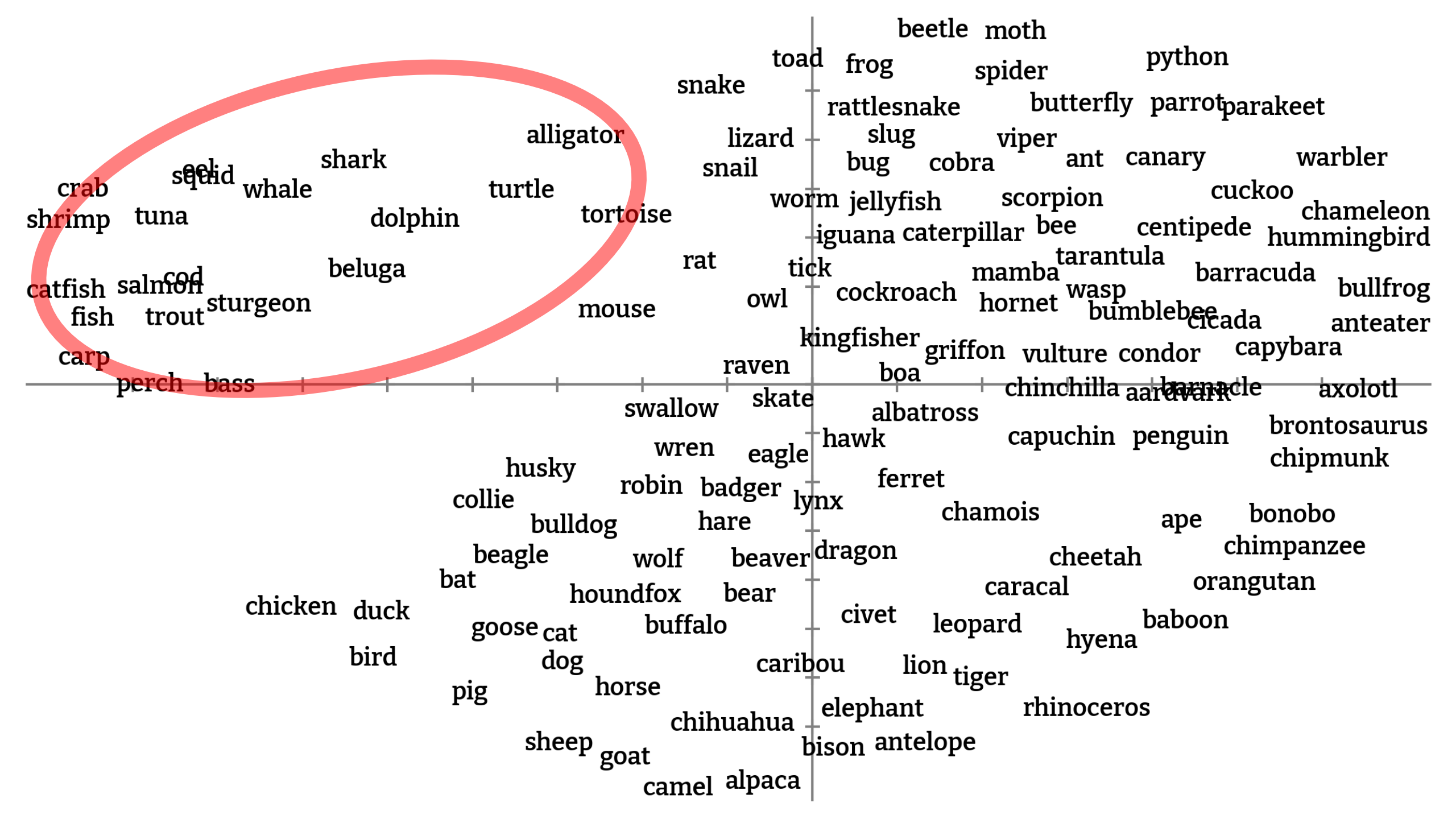
$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + \cdots + a_{1m}x_m \\ a_{21}x_1 + \cdots + a_{2m}x_m \\ \vdots \\ a_{n1}x_1 + \cdots + a_{nm}x_m \end{bmatrix}$$

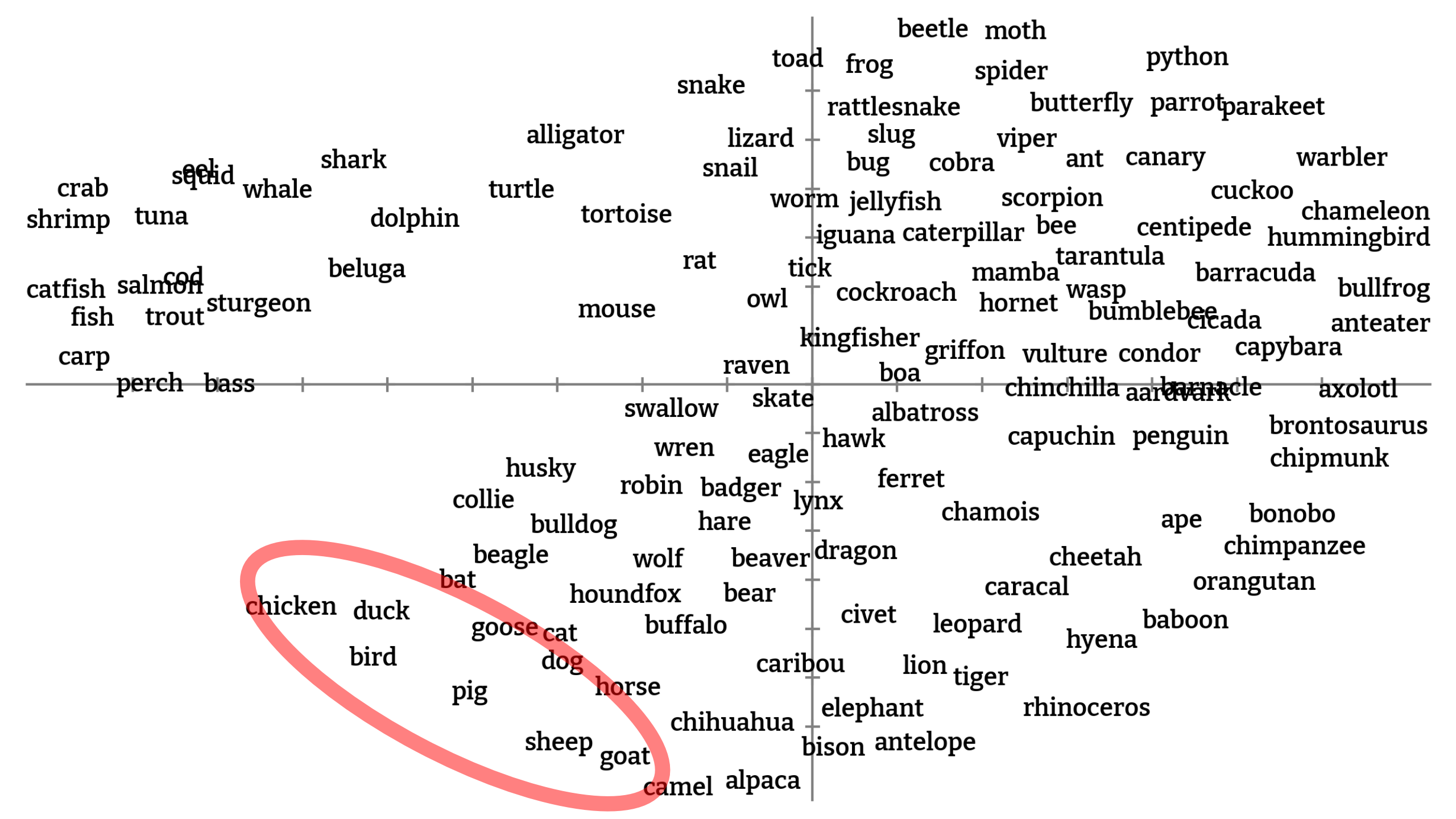
Word Embeddings

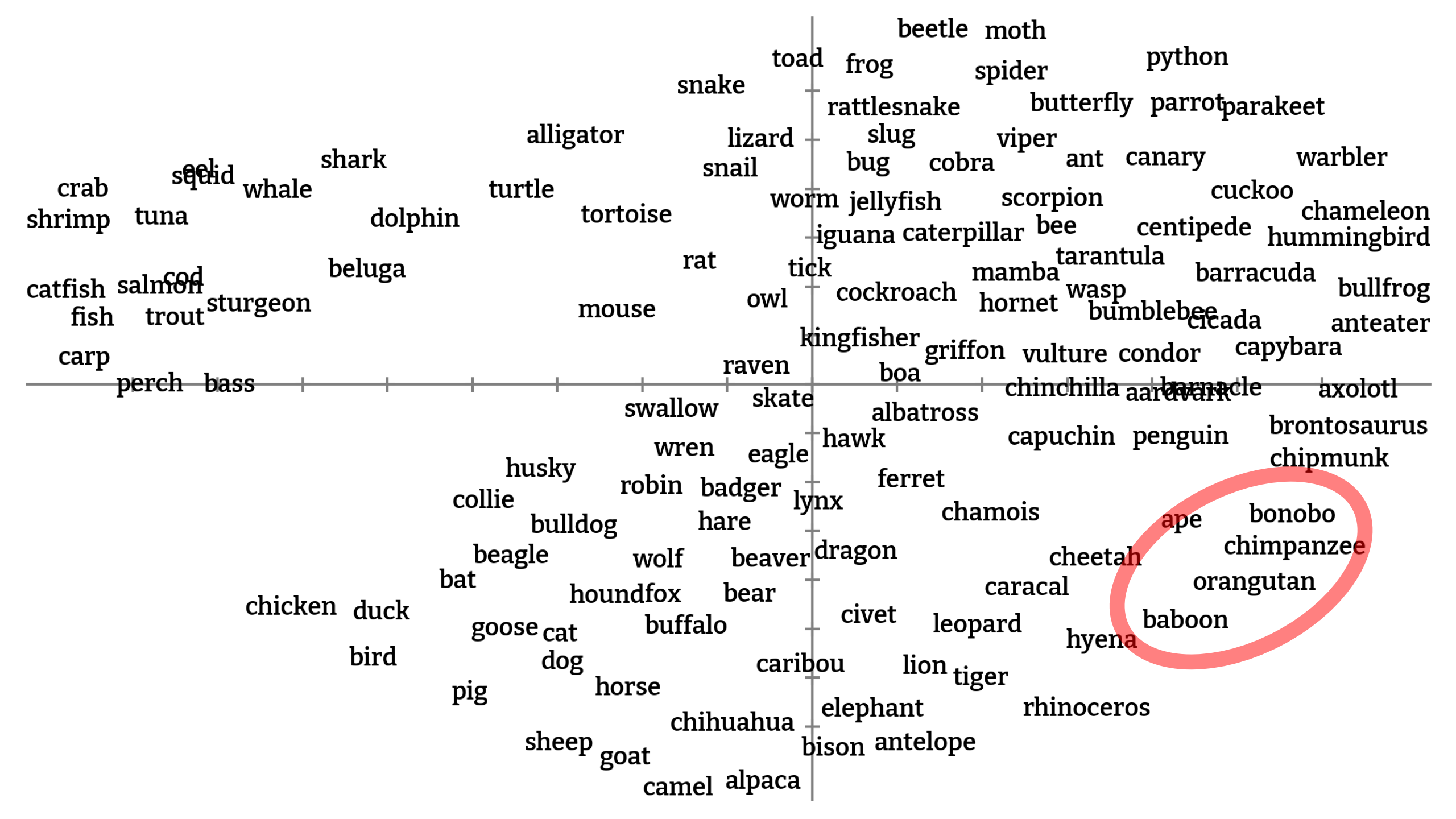

Word
Vector
cat
-0.294
0.332
-0.047
-0.122
0.072
-0.234
-0.062
-0.004
-0.395
-0.694
0.367
...

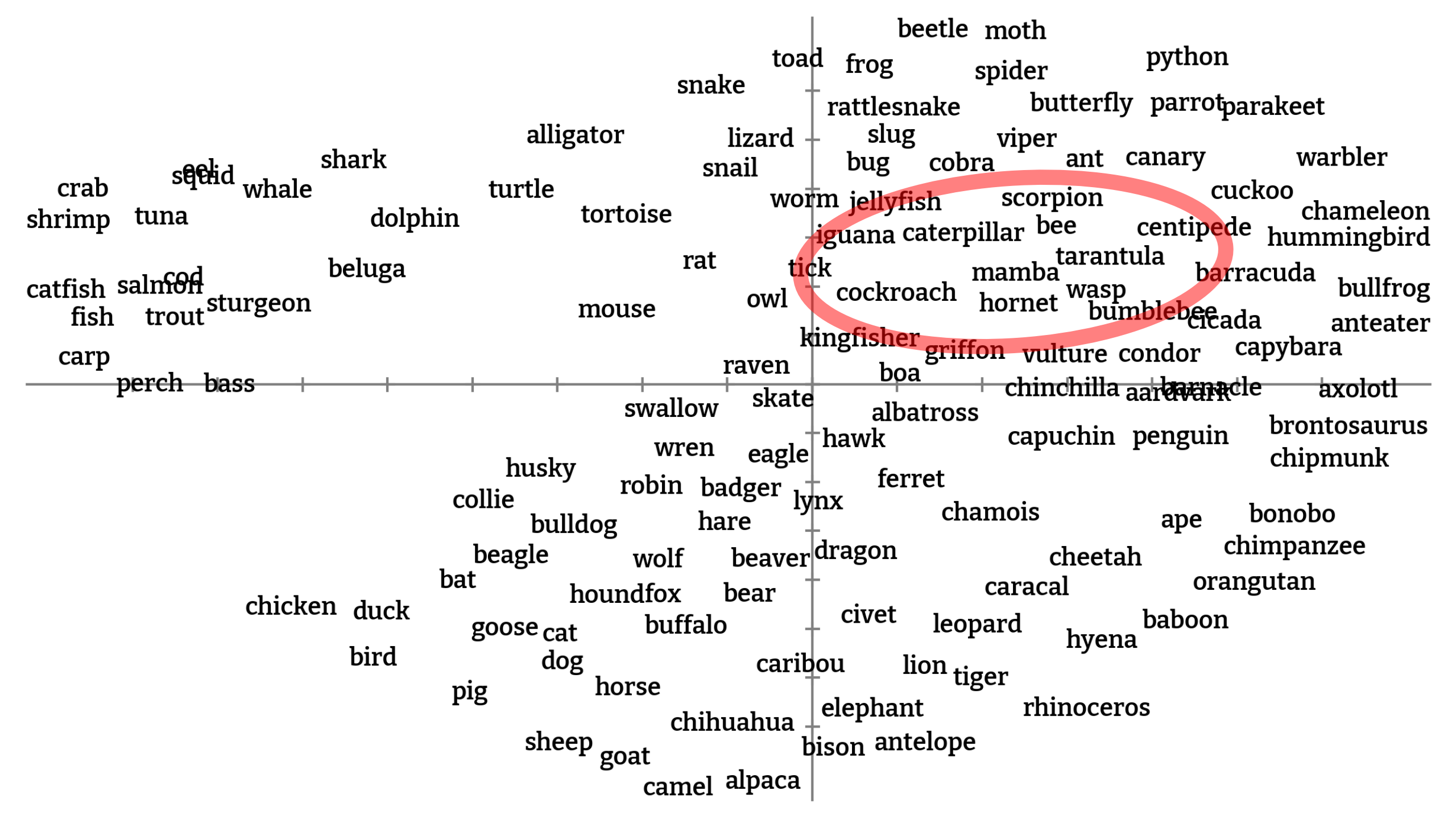
aardvark	-0.349260	-0.553240	-0.430850	0.510470	0.407670	0.095694	-0.176890	0.034768
albatross	0.080120	0.198130	-0.001076	0.418300	0.064855	0.232080	-0.389680	0.288760
alligator	0.648170	0.293300	-0.119370	0.320080	0.122710	-0.216850	-0.082051	-0.169630
alpaca	-0.195220	0.471310	-0.015158	0.180110	0.044267	-0.431260	0.025359	-0.489250
ant	-0.264620	0.803710	0.241130	0.750880	0.139540	0.212790	0.074054	0.503050
anteater	0.055322	-0.532730	-0.130940	-0.170930	0.062775	0.466880	-0.250150	-0.056398
antelope	-0.061041	0.392730	-0.360570	0.070092	0.631310	-0.175620	0.012631	0.254510
ape	-0.412370	0.020455	-0.133590	0.085776	0.058692	0.284700	-0.095120	-0.029808
axolotl	0.460610	-0.021001	0.270570	-0.325080	-0.199290	0.367780	0.098351	0.111730
baboon	-0.511340	0.200730	-0.791900	0.347910	0.320080	0.224010	-0.310820	-0.437500


beetle
moth
toad
frog
spider
python
snake
rattlesnake
butterfly
parrot
parakeet
alligator
lizard
slug
viper
shark
snail
bug
cobra
ant
canary
warbler
eel
squid
crab
whale
turtle
worm
jellyfish
scorpion
cuckoo
shrimp
tuna
dolphin
tortoise
iguana
caterpillar
bee
centipede
chameleon
hummingbird
rat
tick
tarantula
beluga
mamba
barracuda
catfish
salmon
cod
cockroach
wasp
bullfrog
sturgeon
owl
hornet
bumblebee
fish
trout
mouse
cicada
anteater
kingfisher
carp
griffon
vulture
condor
capybara
raven
boa
perch
bass
chinchilla
aardvark
barnacle
axolotl
skate
swallow
albatross
brontosaurus
hawk
capuchin
penguin
wren
eagle
chipmunk
husky
ferret
robin
badger
lynx
collie
chamois
ape
bonobo
bulldog
hare
dragon
chimpanzee
beagle
wolf
beaver
cheetah
bat
caracal
orangutan
houndfox
bear
chicken
duck
civet
leopard
baboon
goose
cat
buffalo
hyena
bird
dog
caribou
lion
tiger
horse
pig
elephant
rhinoceros
chihuahua
sheep
goat
bison
antelope
camel
alpaca


beetle
moth
toad
frog
spider
python
snake
rattlesnake
butterfly
parrot
parakeet
alligator
lizard
slug
viper
shark
snail
bug
cobra
ant
canary
warbler
eel
squid
crab
whale
turtle
worm
jellyfish
scorpion
cuckoo
shrimp
tuna
dolphin
tortoise
iguana
caterpillar
bee
centipede
chameleon
hummingbird
rat
tick
tarantula
beluga
mamba
barracuda
catfish
salmon
cod
owl
cockroach
wasp
hornet
bullfrog
sturgeon
mouse
bumblebee
cicada
anteater
fish
trout
kingfisher
griffon
vulture
condor
capybara
carp
raven
boa
chinchilla
aardvark
barnacle
axolotl
perch
bass
skate
swallow
albatross
brontosaurus
hawk
capuchin
penguin
wren
eagle
chipmunk
husky
ferret
robin
badger
lynx
collie
chamois
ape
bonobo
bulldog
hare
dragon
beagle
wolf
beaver
cheetah
chimpanzee
bat
caracal
orangutan
houndfox
bear
chicken
duck
civet
leopard
baboon
goose
cat
buffalo
hyena
bird
dog
caribou
lion
tiger
horse
pig
elephant
rhinoceros
chihuahua
sheep
goat
bison
antelope
camel
alpaca


beetle
moth
toad
frog
spider
python
snake
rattlesnake
butterfly
parrot
parakeet
alligator
lizard
slug
viper
shark
snail
bug
cobra
ant
canary
warbler
eel
squid
crab
whale
turtle
worm
jellyfish
scorpion
cuckoo
shrimp
tuna
dolphin
tortoise
chameleon
iguana
caterpillar
bee
centipede
hummingbird
rat
tarantula
beluga
tick
mamba
barracuda
catfish
salmon
cod
cockroach
wasp
bullfrog
sturgeon
owl
hornet
fish
trout
mouse
bumblebee
cicada
anteater
kingfisher
carp
griffon
vulture
condor
capybara
raven
boa
perch
bass
chinchilla
aardvark
barnacle
axolotl
skate
swallow
albatross
brontosaurus
hawk
capuchin
penguin
wren
eagle
chipmunk
husky
ferret
robin
badger
lynx
collie
chamois
ape
bonobo
bulldog
hare
dragon
chimpanzee
beagle
wolf
beaver
cheetah
bat
caracal
orangutan
houndfox
bear
chicken
duck
civet
leopard
baboon
goose
cat
buffalo
hyena
bird
caribou
lion
dog
tiger
pig
horse
elephant
rhinoceros
chihuahua
sheep
bison
antelope
goat
camel
alpaca


beetle
moth
toad
frog
spider
python
snake
rattlesnake
butterfly
parrot
parakeet
alligator
lizard
slug
viper
shark
snail
bug
cobra
ant
canary
warbler
eel
squid
crab
whale
turtle
worm
jellyfish
scorpion
cuckoo
shrimp
tuna
dolphin
tortoise
iguana
caterpillar
bee
centipede
chameleon
hummingbird
rat
tarantula
beluga
tick
mamba
barracuda
cod
catfish
salmon
cockroach
wasp
bullfrog
sturgeon
owl
hornet
fish
trout
mouse
bumblebee
cicada
anteater
kingfisher
carp
griffon
vulture
condor
capybara
raven
boa
perch
bass
chinchilla
aardvark
barnacle
axolotl
skate
swallow
albatross
brontosaurus
hawk
capuchin
penguin
wren
eagle
chipmunk
husky
ferret
robin
badger
lynx
collie
chamois
ape
bonobo
bulldog
hare
dragon
chimpanzee
beagle
wolf
beaver
cheetah
bat
caracal
orangutan
houndfox
bear
chicken
duck
civet
baboon
goose
cat
buffalo
leopard
hyena
bird
dog
caribou
lion
tiger
horse
pig
elephant
rhinoceros
chihuahua
sheep
bison
antelope
goat
camel
alpaca


beetle
moth
toad
frog
spider
python
snake
rattlesnake
butterfly
parrot
parakeet
alligator
lizard
slug
viper
shark
snail
bug
cobra
ant
canary
warbler
crab
whale
turtle
worm
jellyfish
scorpion
cuckoo
shrimp
tuna
dolphin
tortoise
chameleon
iguana
caterpillar
bee
centipede
hummingbird
rat
tarantula
beluga
tick
mamba
barracuda
catfish
salmon
cod
cockroach
wasp
bullfrog
fish
trout
sturgeon
mouse
owl
hornet
bumblebee
cicada
anteater
kingfisher
carp
griffon
vulture
condor
capybara
raven
boa
perch
bass
chinchilla
aardvark
barnacle
axolotl
skate
swallow
albatross
brontosaurus
hawk
capuchin
penguin
wren
eagle
chipmunk
husky
ferret
robin
badger
lynx
collie
chamois
ape
bonobo
bulldog
hare
dragon
chimpanzee
beagle
wolf
beaver
cheetah
bat
caracal
orangutan
houndfox
bear
chicken
duck
civet
baboon
goose
cat
buffalo
leopard
hyena
bird
dog
caribou
lion
tiger
pig
horse
elephant
rhinoceros
chihuahua
sheep
bison
antelope
goat
camel
alpaca

Text Embeddings

Text → Vector

Alice was beginning to get very tired of sitting by her sister on the bank, and of having nothing to do: once or twice she had peeped into the book her sister was reading, but it had no pictures or conversations in it, "and what is the use of a book," thought Alice "without pictures or conversations?"

So she was considering in her own mind (as well as she could, for the hot day made her feel very sleepy and stupid), whether the pleasure of making a daisy-chain would be worth the trouble of getting up and picking the daisies, when suddenly a White Rabbit with pink eyes ran close by her.

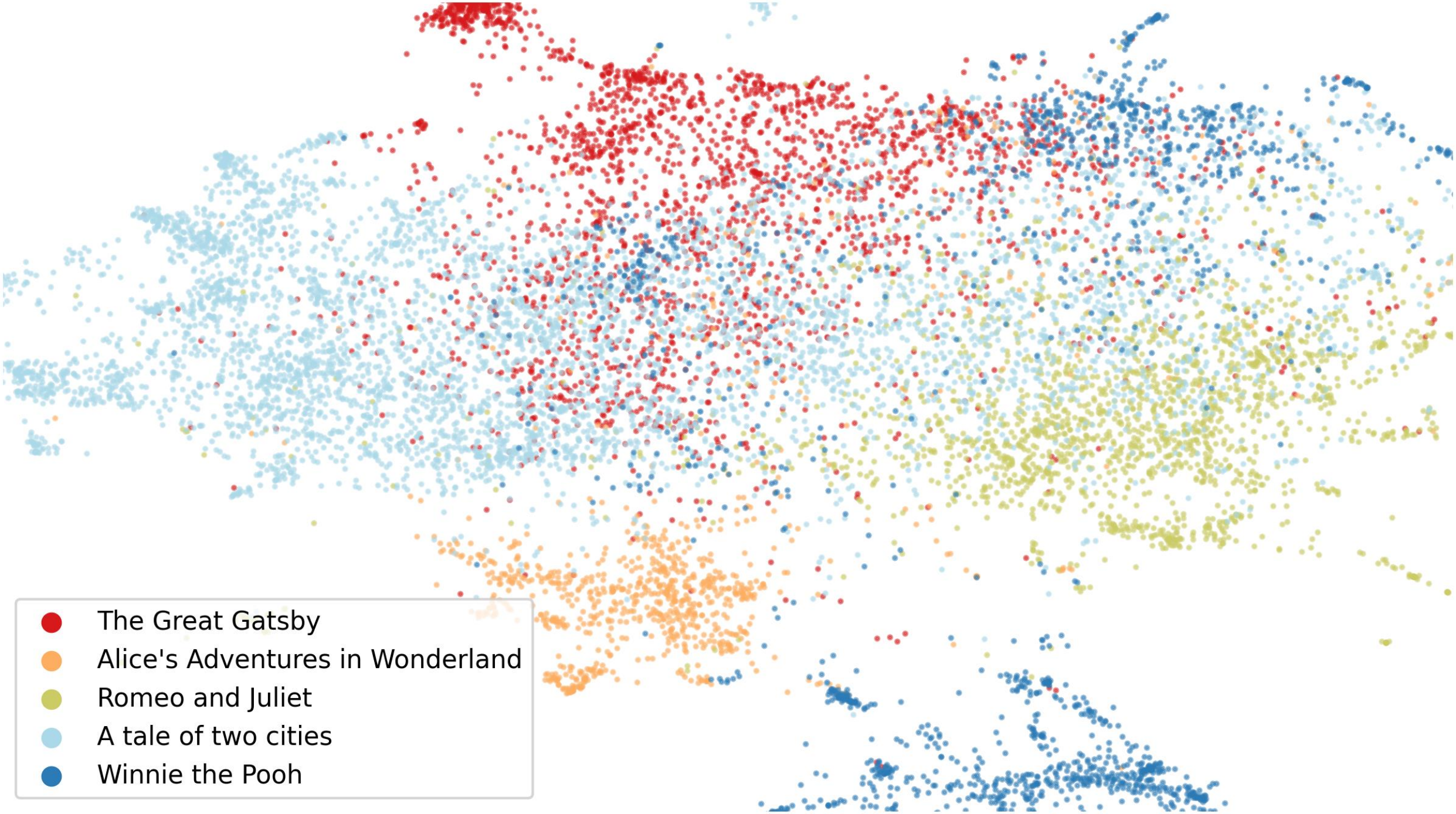
There was nothing so _very_ remarkable in that; nor did Alice think it so _very_ much out of the way to hear the Rabbit say to itself, "Oh dear! Oh dear! I shall be late!" (when she thought it over afterwards, it occurred to her that she ought to have wondered at this, but at the time it all seemed quite natural); but when the Rabbit actually _took a watch out of its waistcoat-pocket_, and looked at it, and then hurried on, Alice started to her feet, for it flashed across her mind that she had never before seen a rabbit with either a waistcoat-pocket, or a watch to take out of it, and burning with curiosity, she ran across the field after it, and fortunately was just in time to see it pop down a large rabbit-hole under the hedge.

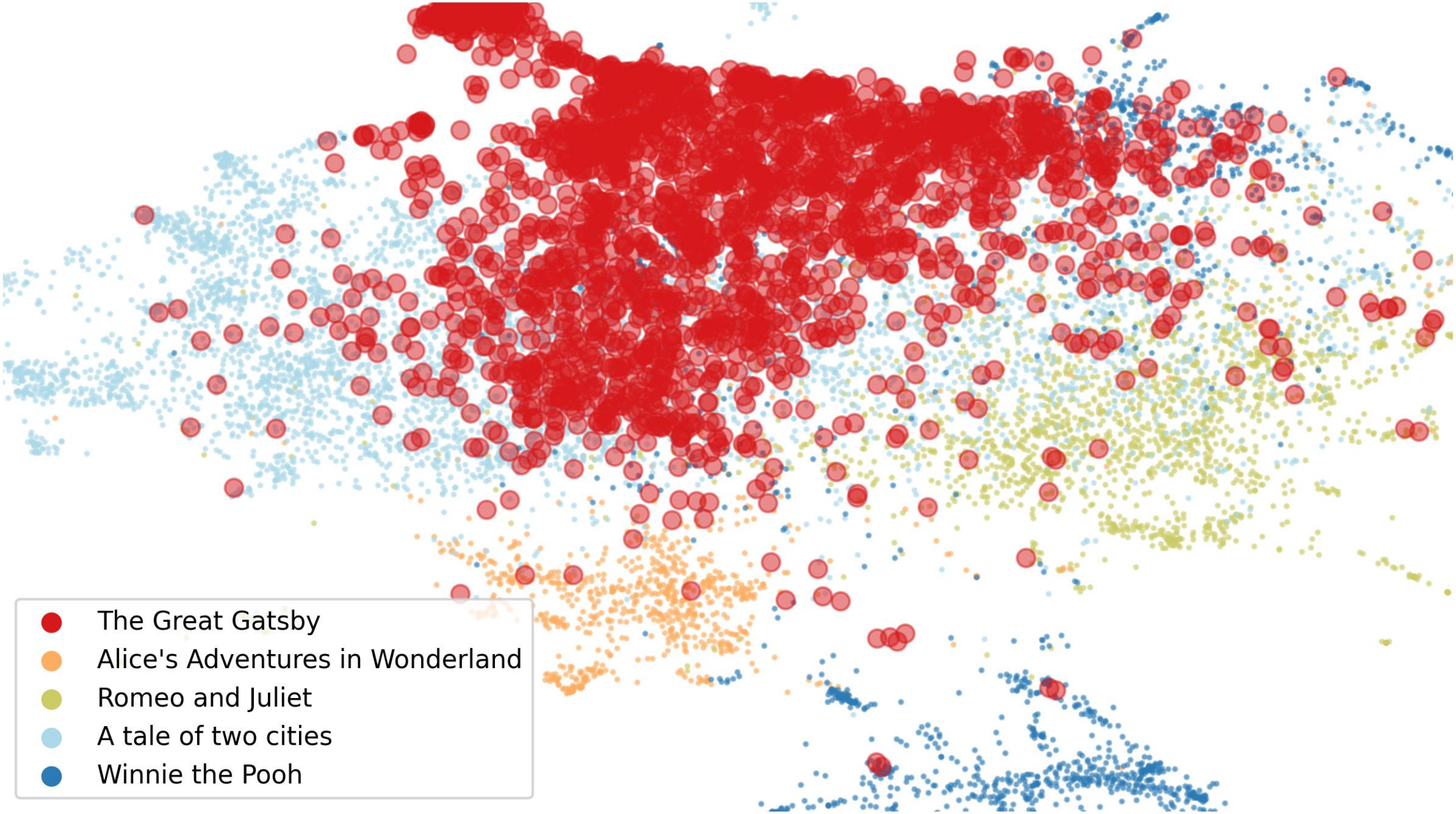
In another moment down went Alice after it, never once considering how in the world she was to get out again.

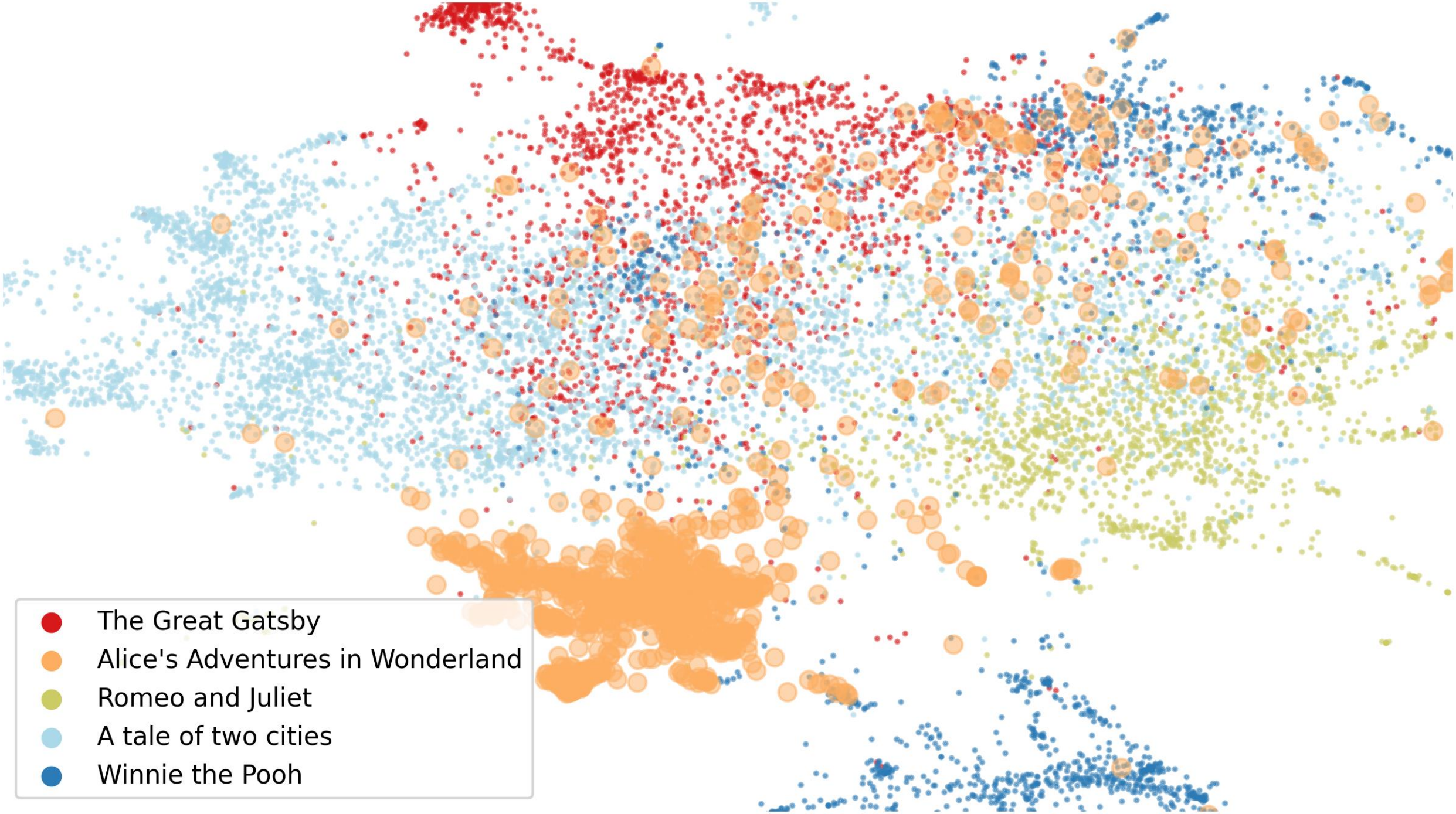
The rabbit-hole went straight on like a tunnel for some way, and then dipped suddenly down, so suddenly that Alice had not a moment to think about stopping herself before she found herself falling down a very deep well.

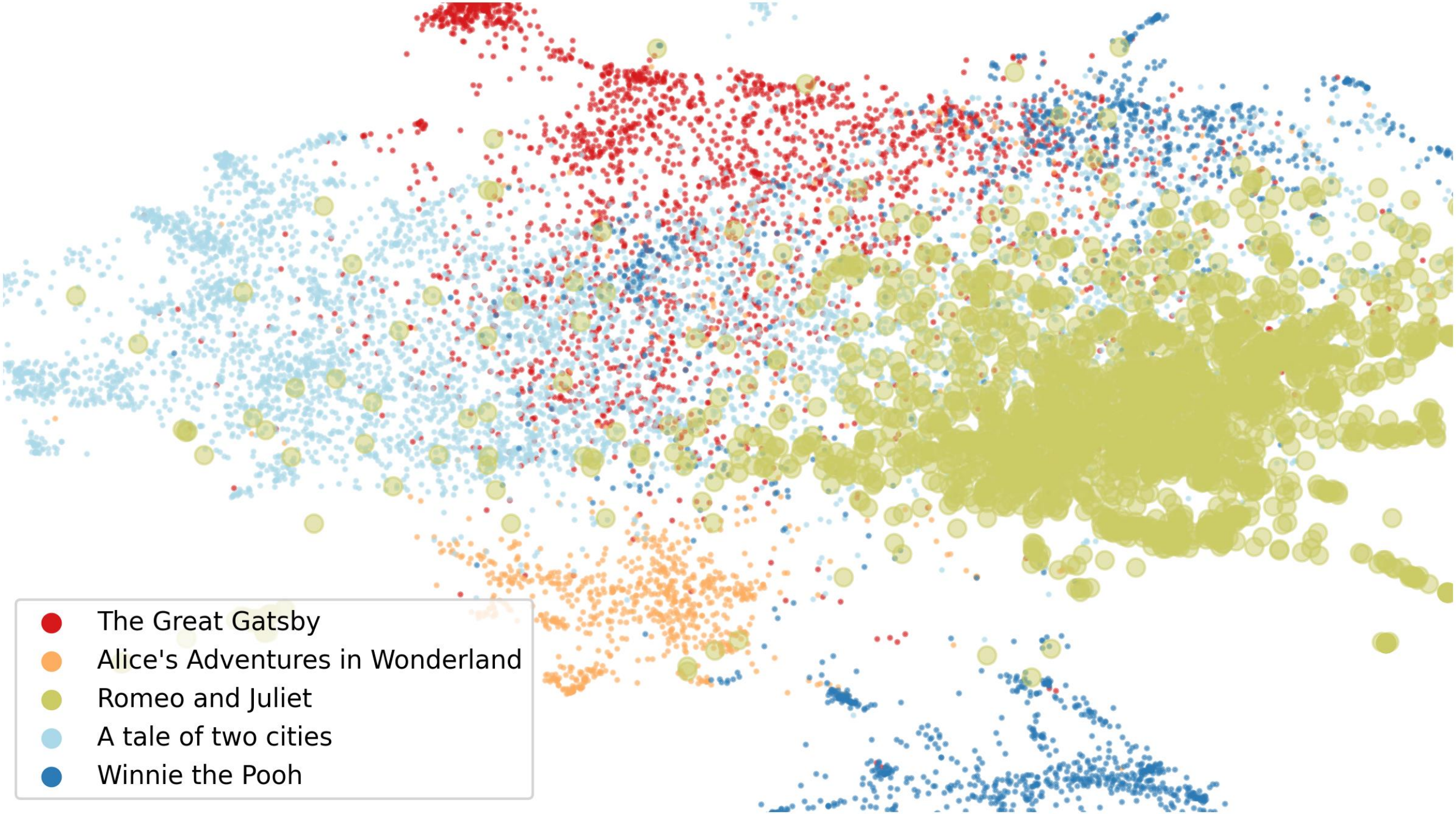
Either the well was very deep, or she fell very slowly, for she had plenty of time as she went down to look about her and to wonder what was going to happen next. First, she tried to look down and make out what she was coming to, but it was too dark to see anything; then she looked at the sides of the well, and noticed that they were filled with cupboards and book-shelves; here and there she saw maps and pictures hung upon pegs. She took down a jar from one of the shelves as she passed; it was labelled "ORANGE MARMALADE", but to her great disappointment it was empty: she did not like to drop the jar for fear of killing somebody underneath, so managed to put it into one of the cupboards as she fell past it.

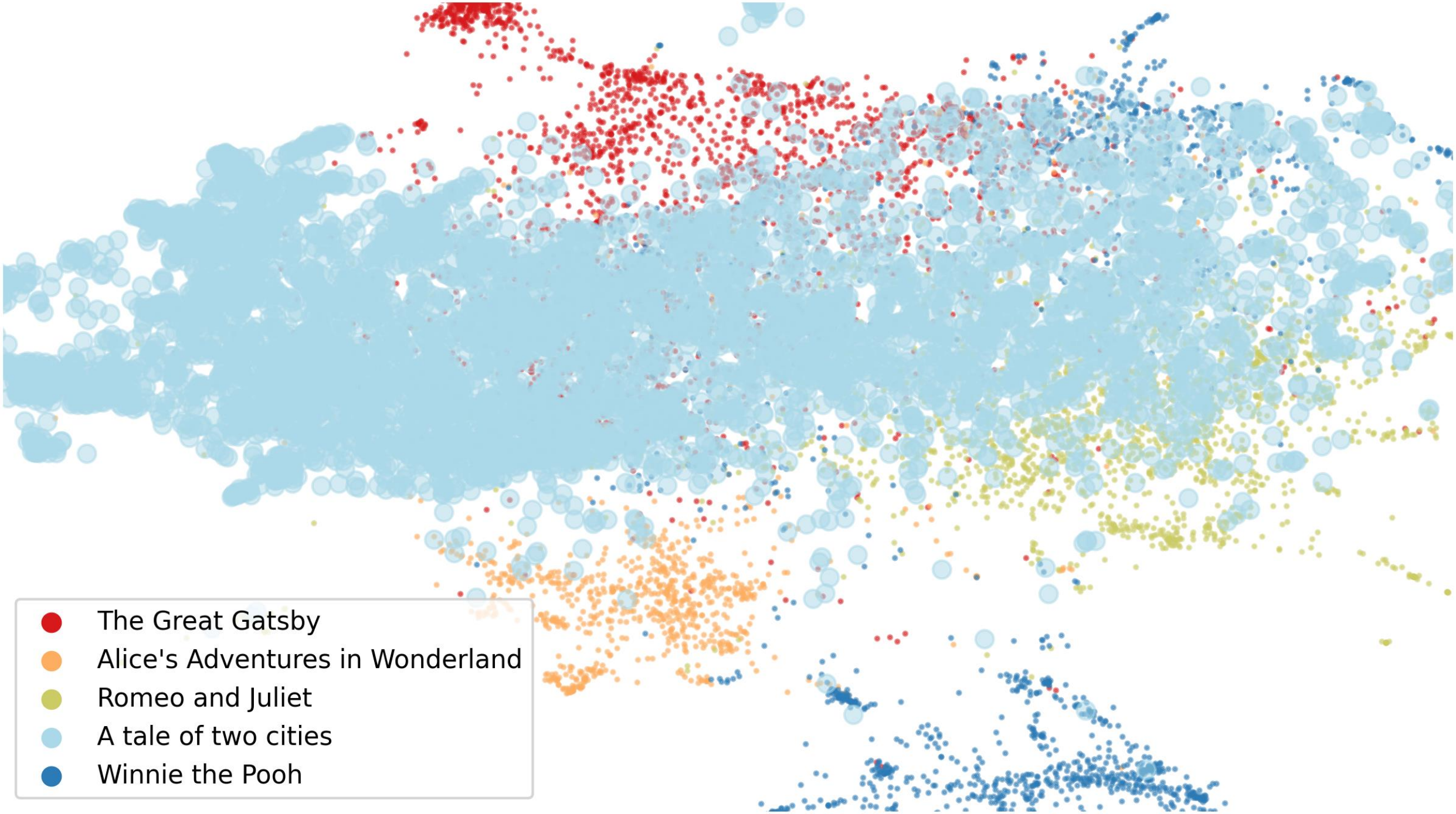
```
 0.066
-0.019
-0.149
 0.283
-0.450
 0.351
-0.411
 0.008
-0.061
-0.399
 0.391
  ...
```

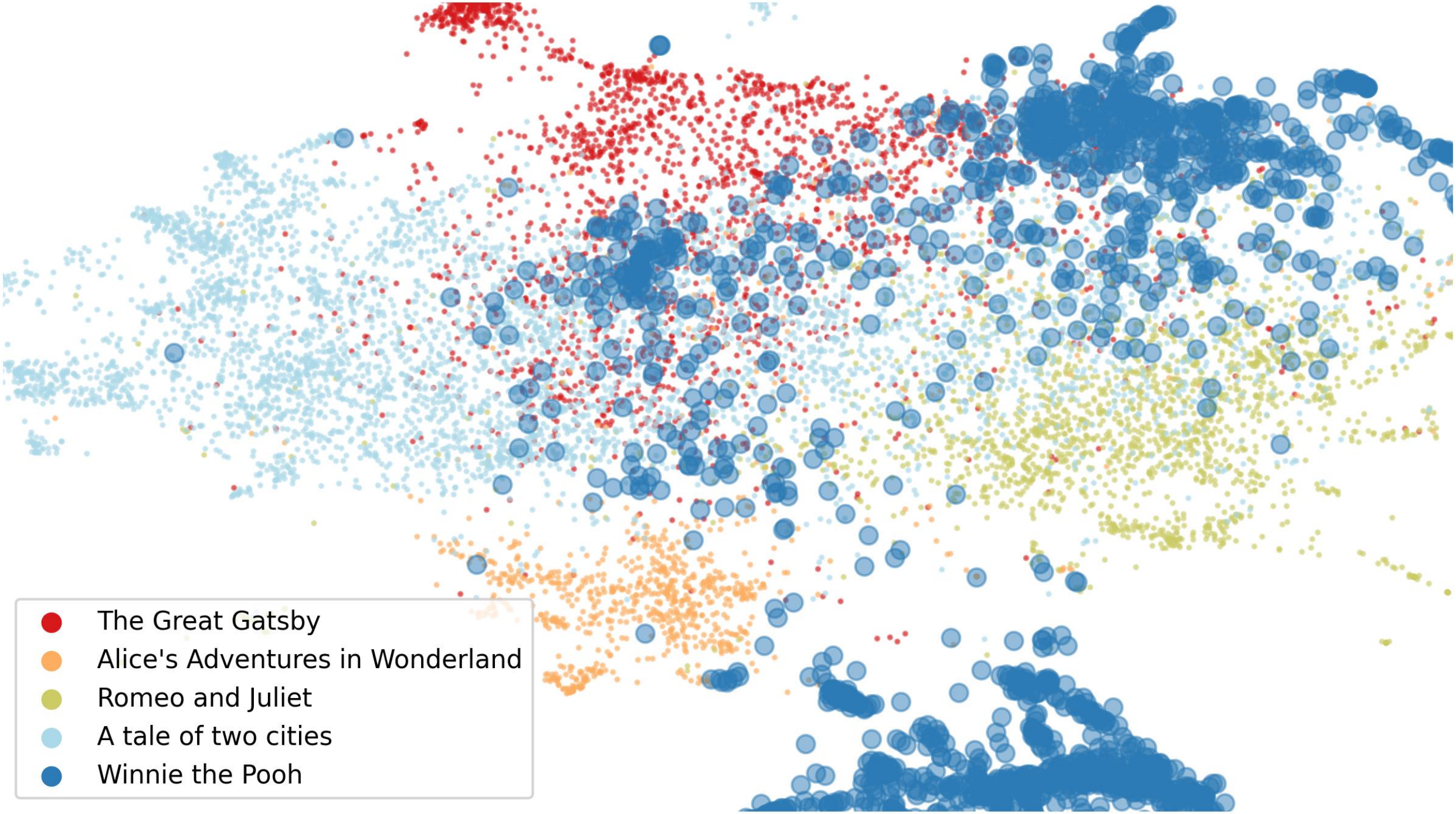

The Great Gatsby
Alice's Adventures in Wonderland
Romeo and Juliet
A tale of two cities
Winnie the Pooh


The Great Gatsby
Alice's Adventures in Wonderland
Romeo and Juliet
A tale of two cities
Winnie the Pooh


The Great Gatsby
Alice's Adventures in Wonderland
Romeo and Juliet
A tale of two cities
Winnie the Pooh

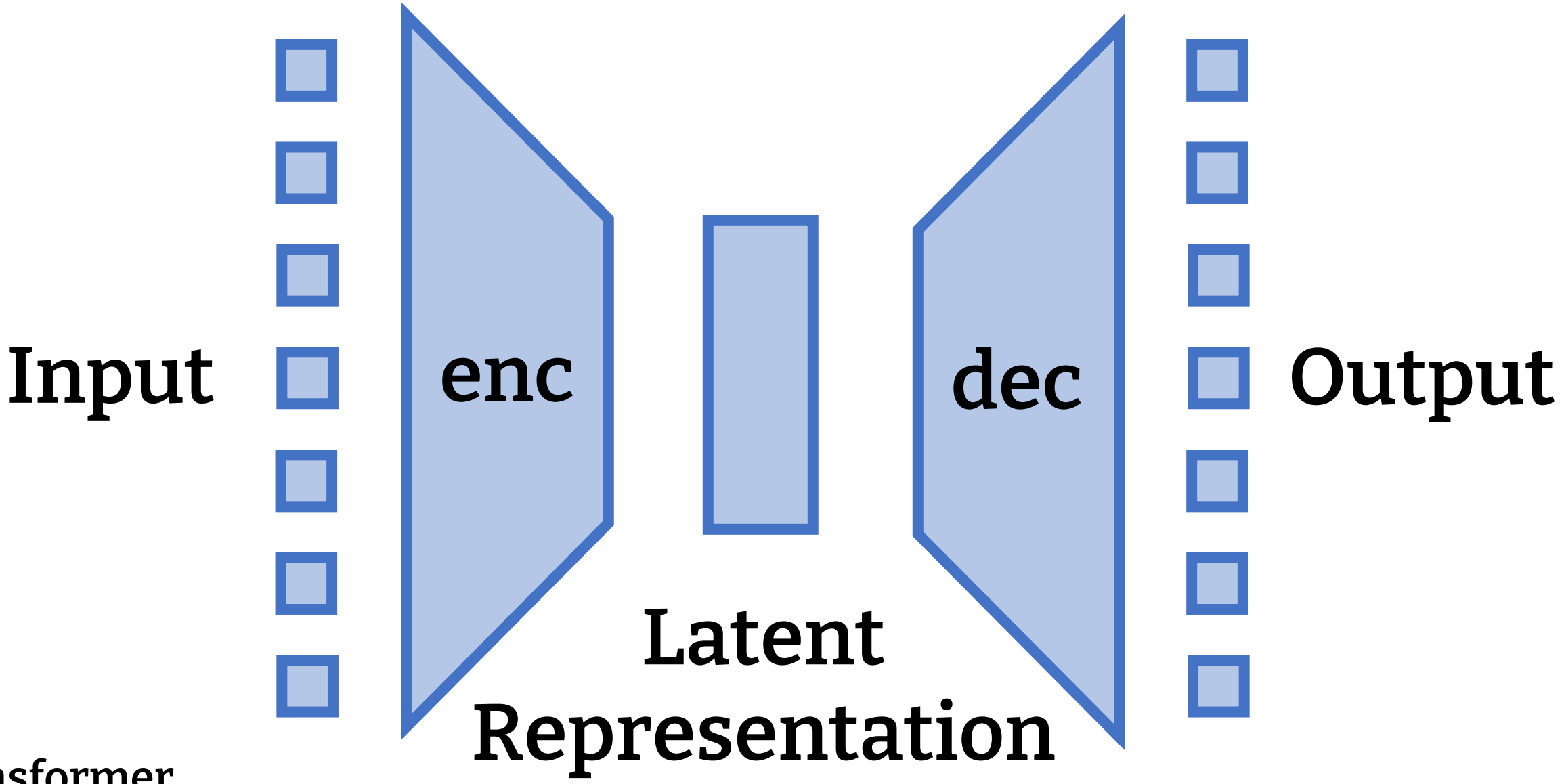

The Great Gatsby
Alice's Adventures in Wonderland
Romeo and Juliet
A tale of two cities
Winnie the Pooh


The Great Gatsby
Alice's Adventures in Wonderland
Romeo and Juliet
A tale of two cities
Winnie the Pooh

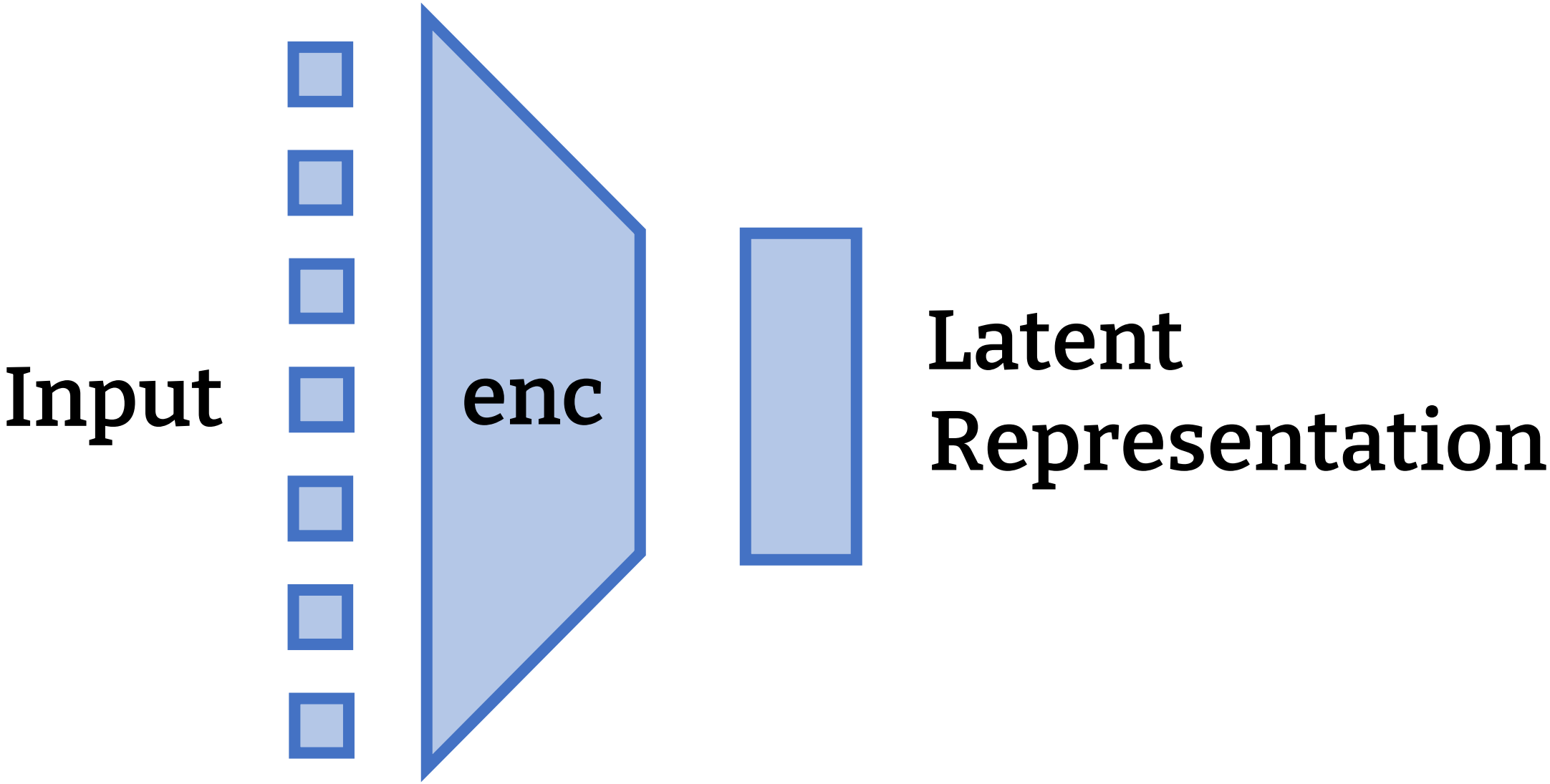

The Great Gatsby
Alice's Adventures in Wonderland
Romeo and Juliet
A tale of two cities
Winnie the Pooh

Attention

Encoder-Decoder

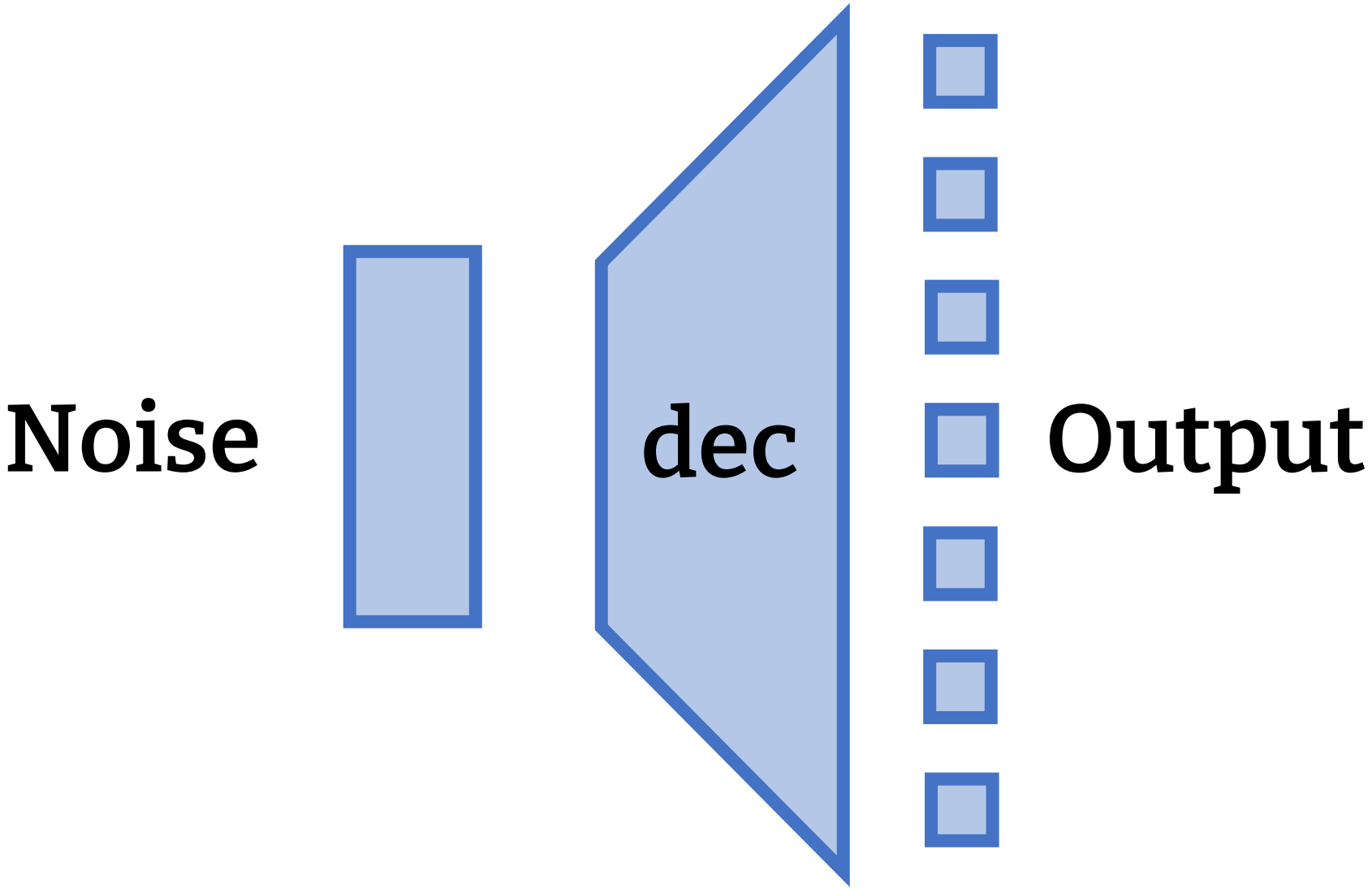


Encoder-only



BERT

Decoder-only



GPT

Differentiable hash table

q: query, row vector
K: keys, one per row
V: values, one per row

$$\mathrm{Att}(q, K, V) = \mathrm{softmax}(q K') V$$

Attention

Q: queries, one per row
K: keys, one per row
V: values, one per row

$$\text{Att}(Q, K, V) = \text{softmax}(Q K') V$$

Attention

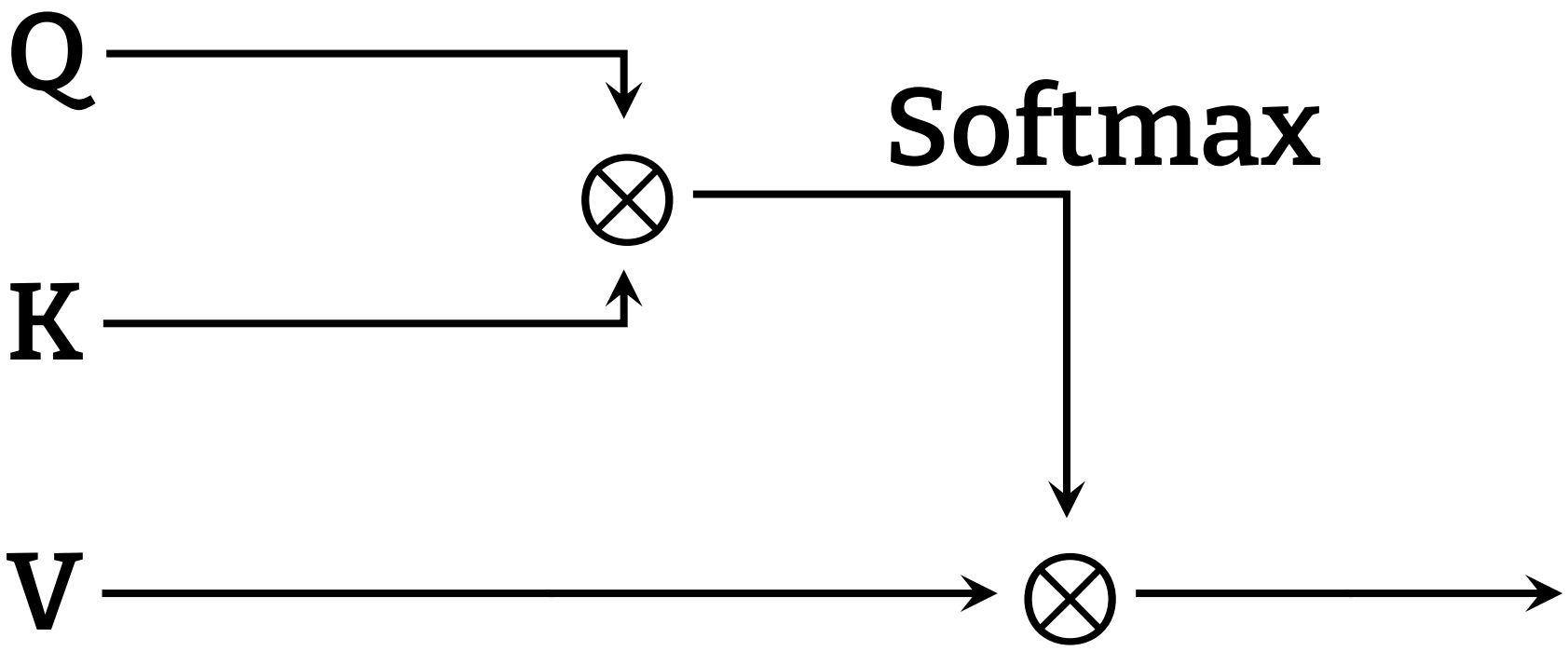
Q: queries, one per row

K: keys, one per row

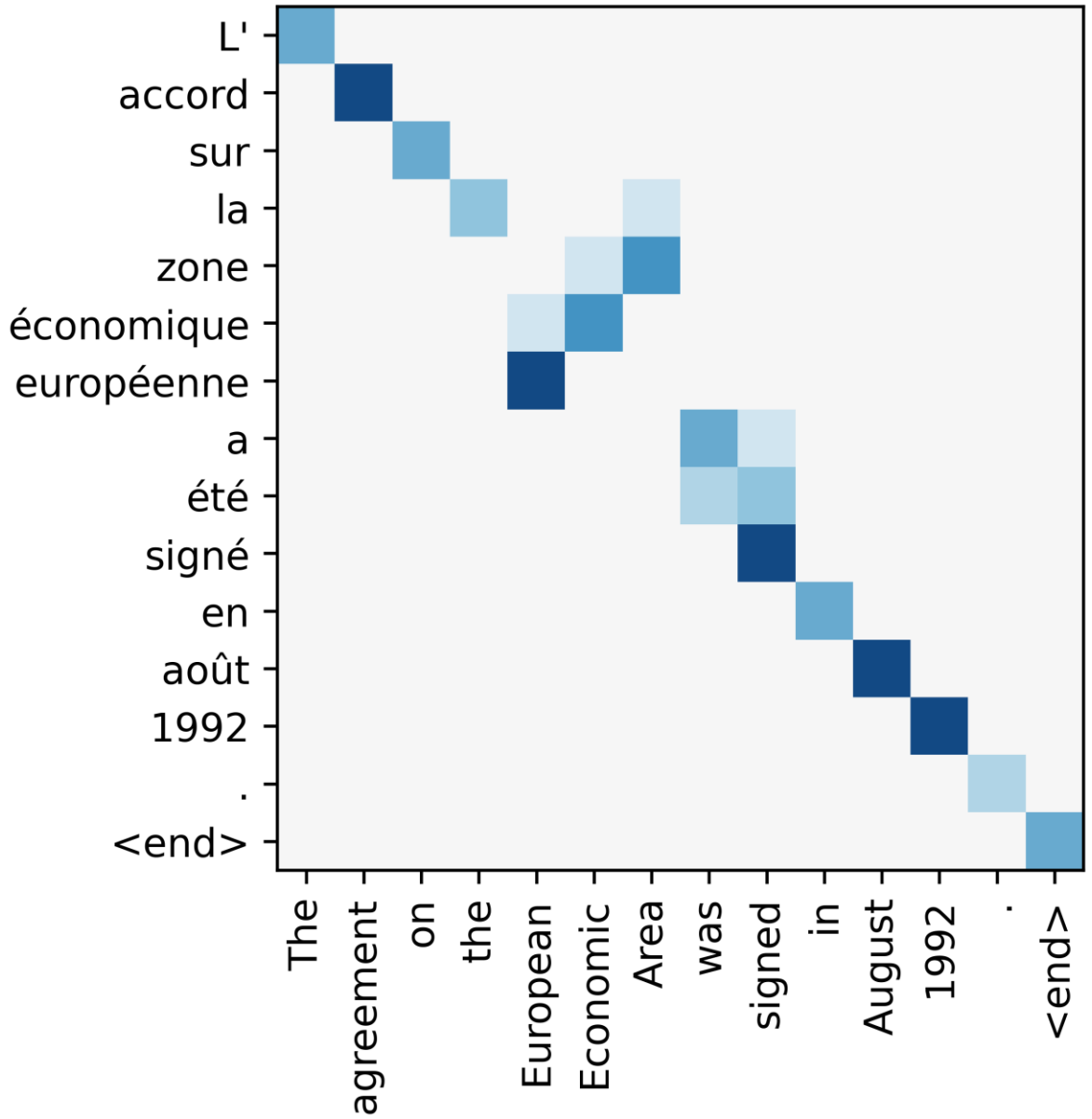
V: values, one per row

$$\text{Att}(Q, K, V) = \text{softmax}(\frac{Q K'}{\sqrt{d}}) V$$

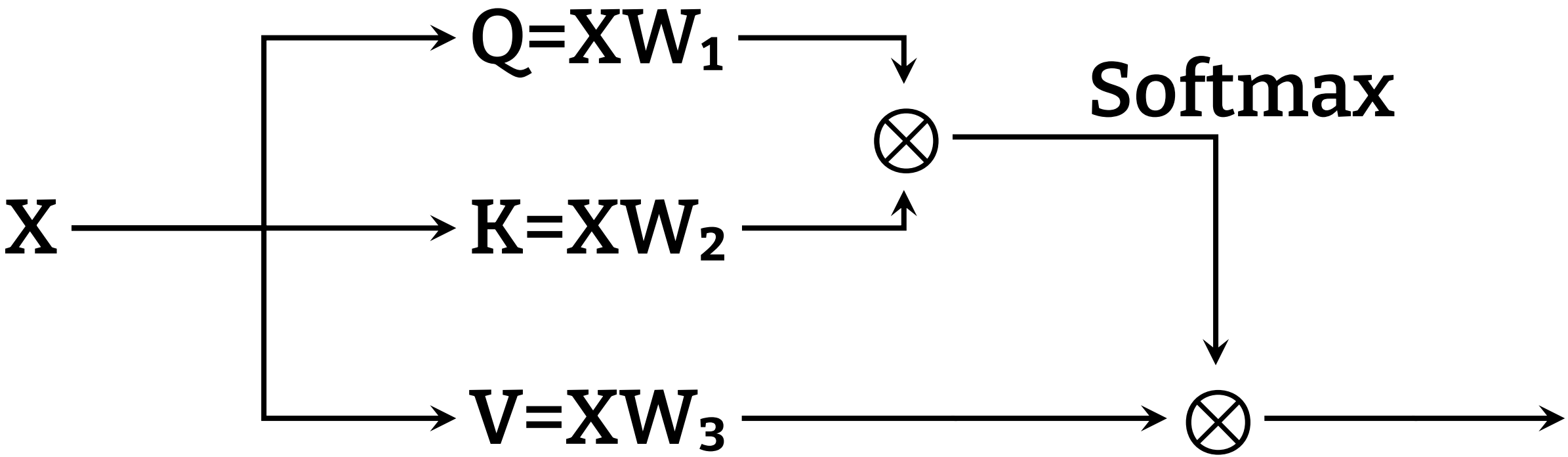
Attention



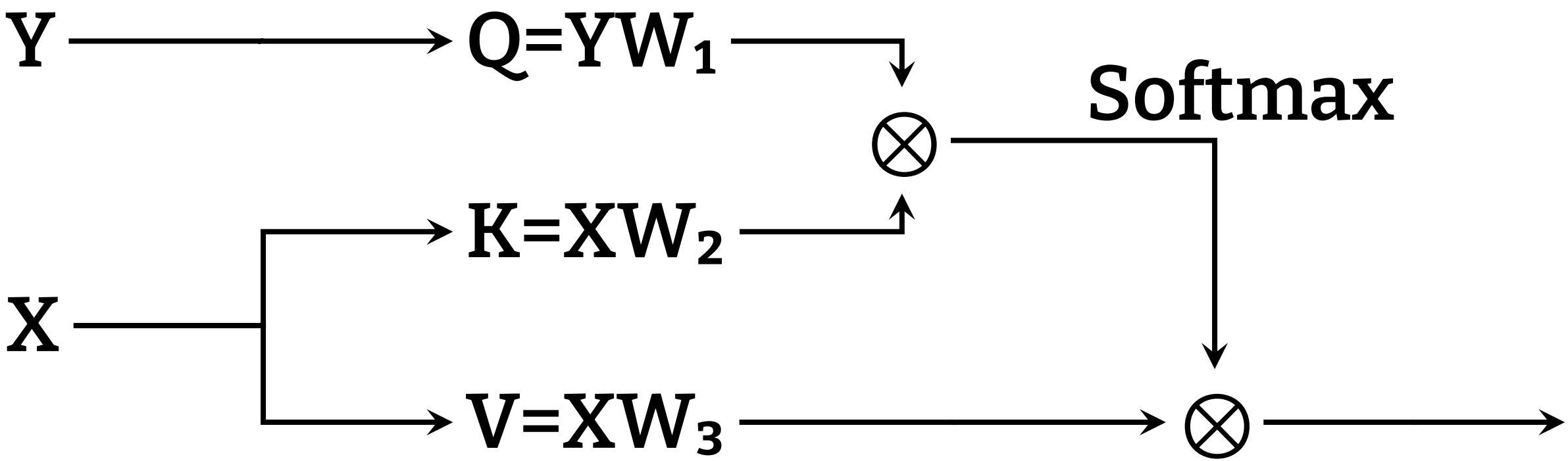
Attention Matrix



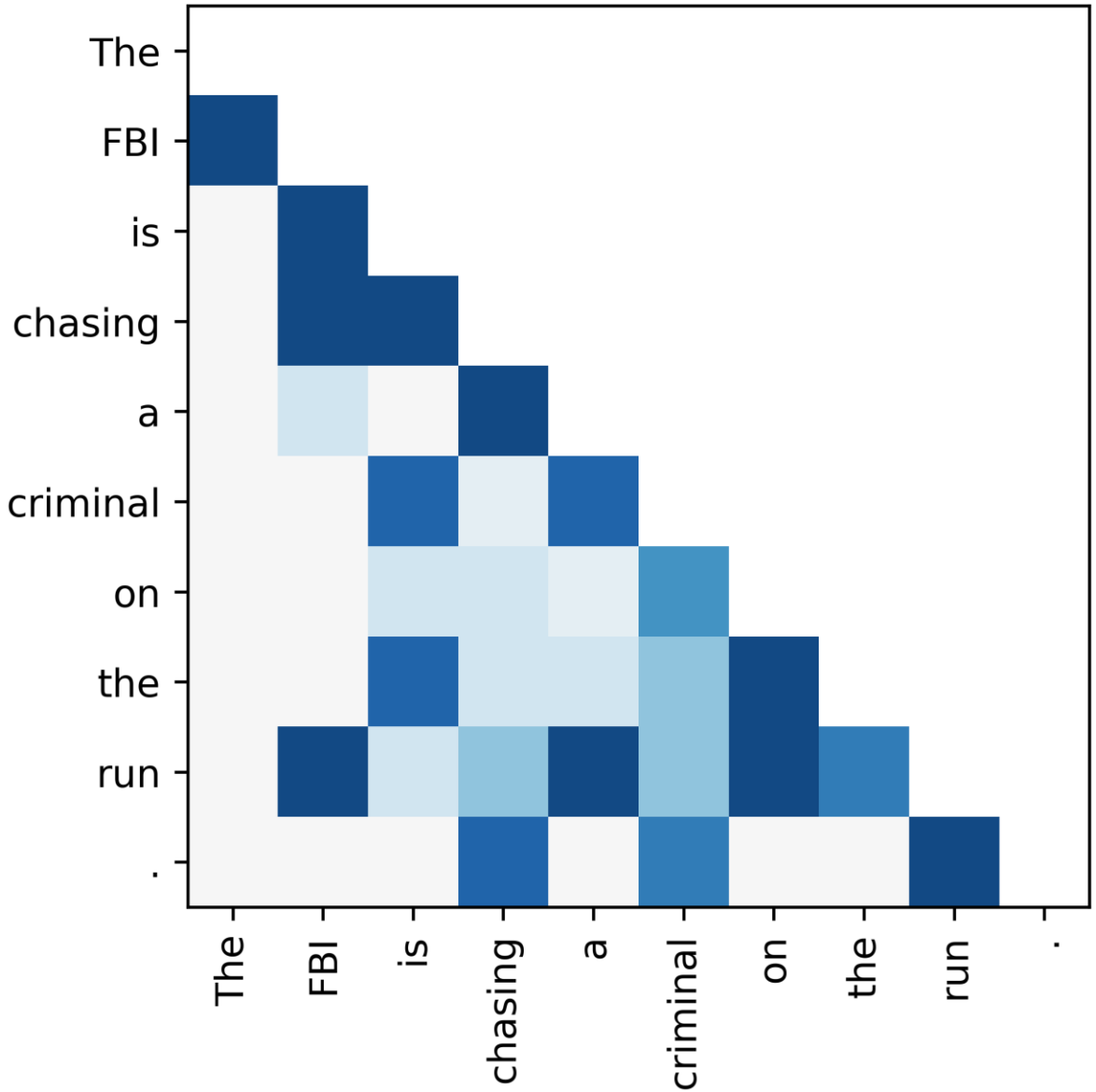
Self-Attention



Cross-Attention

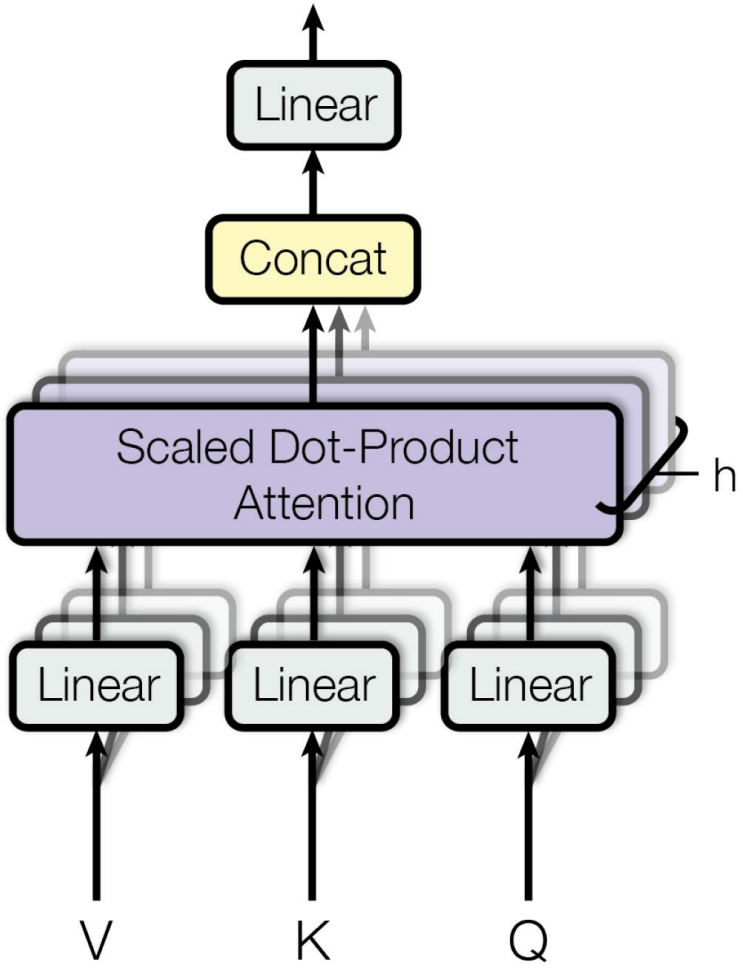


Masked Self-Attention



Multi-Head Self-Attention

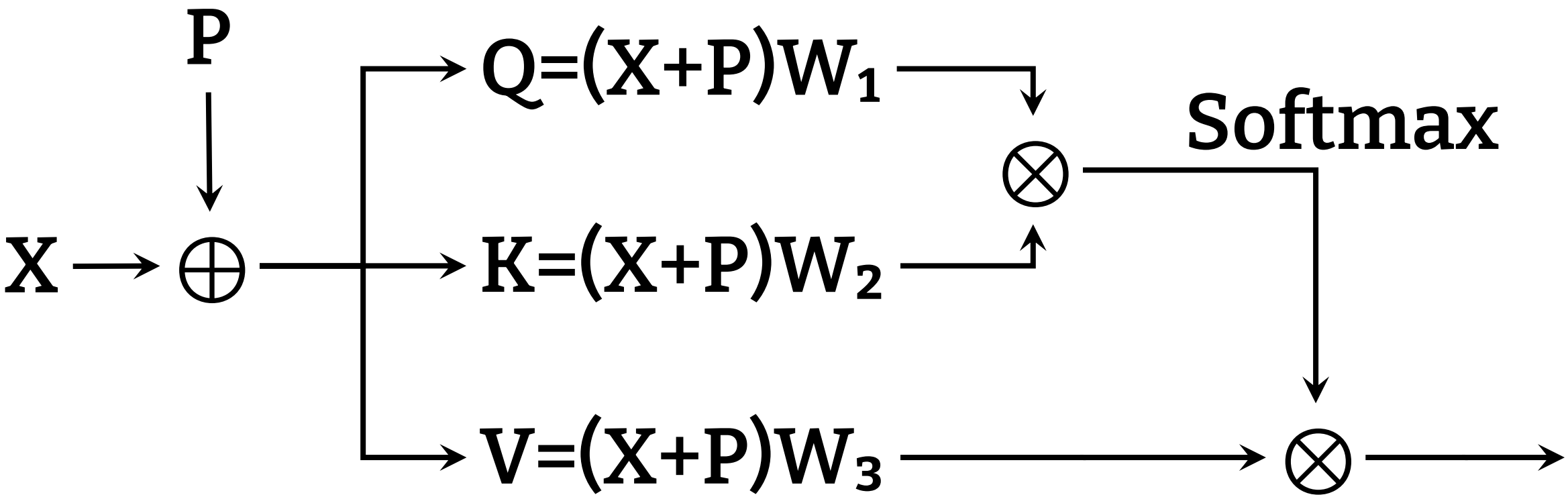
- Several self-attention blocks in parallel



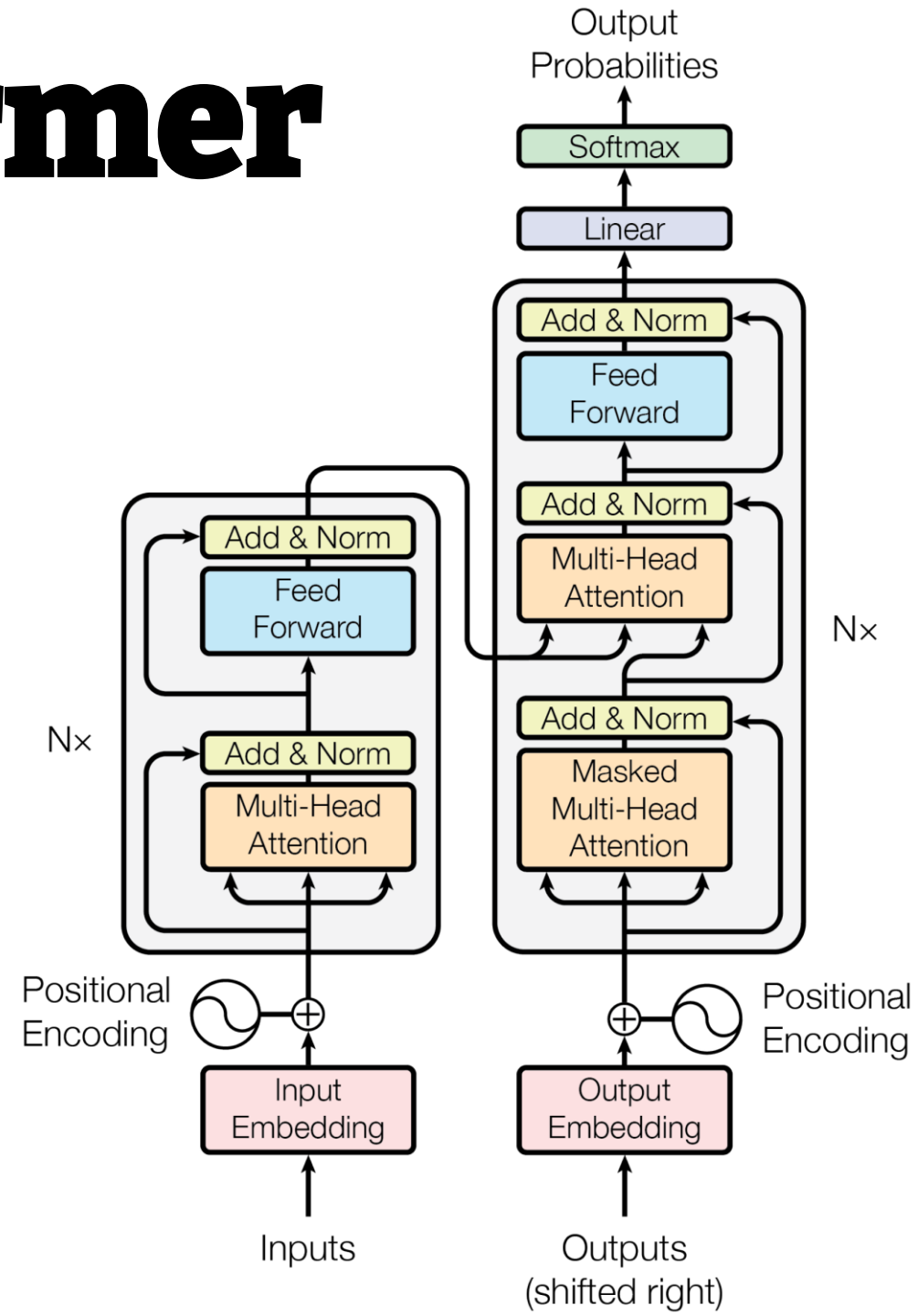
Positional Encoding

- Attention is permutation-invariant
- Add $\sin(k \cdot i \cdot 2\pi/N)$ and $\cos(k \cdot i \cdot 2\pi/N)$ as features, for token $i \in [\![1, N]\!]$, for a few values of k
- Often, just added to X

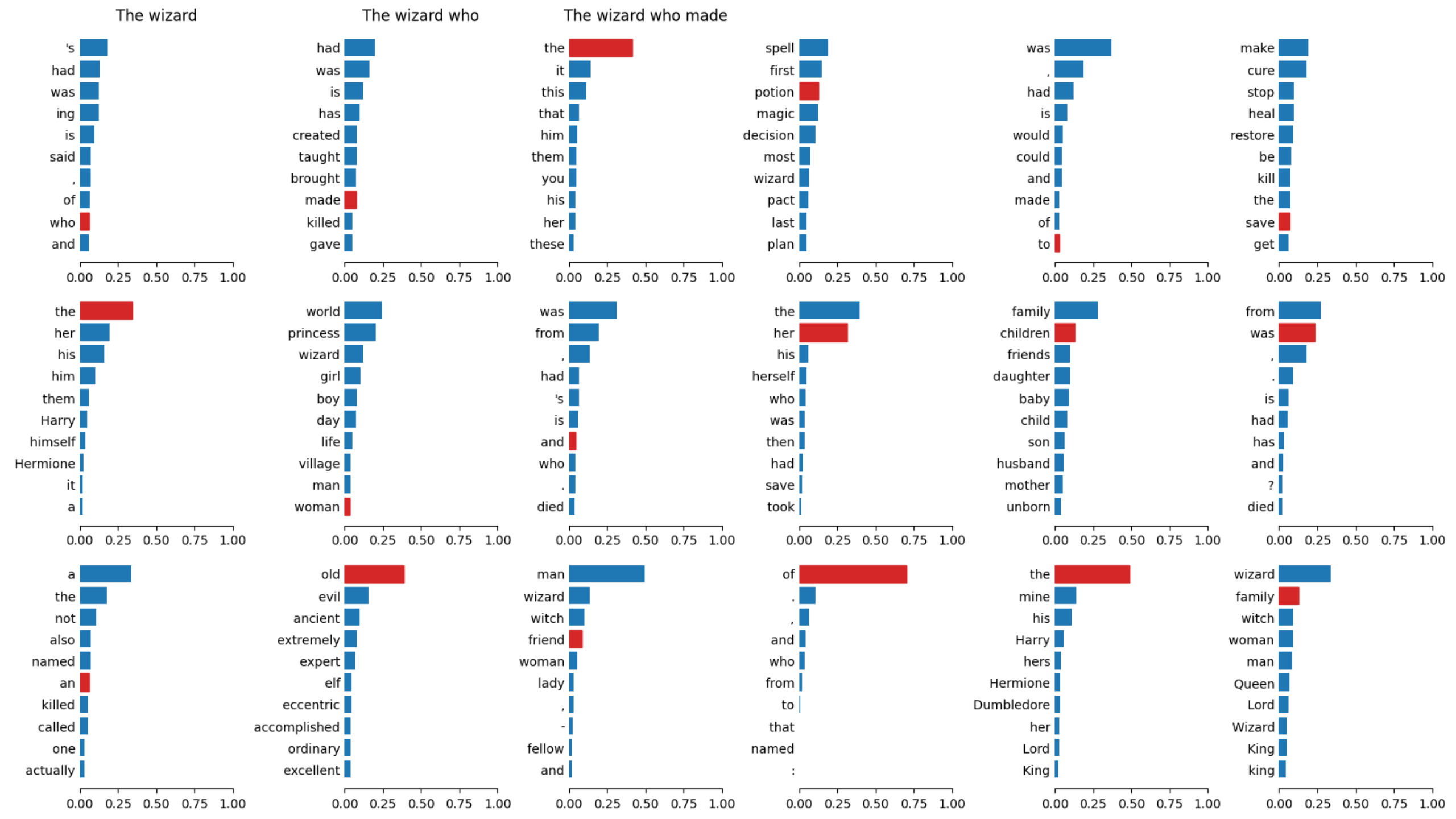
Positional Encoding



Transformer

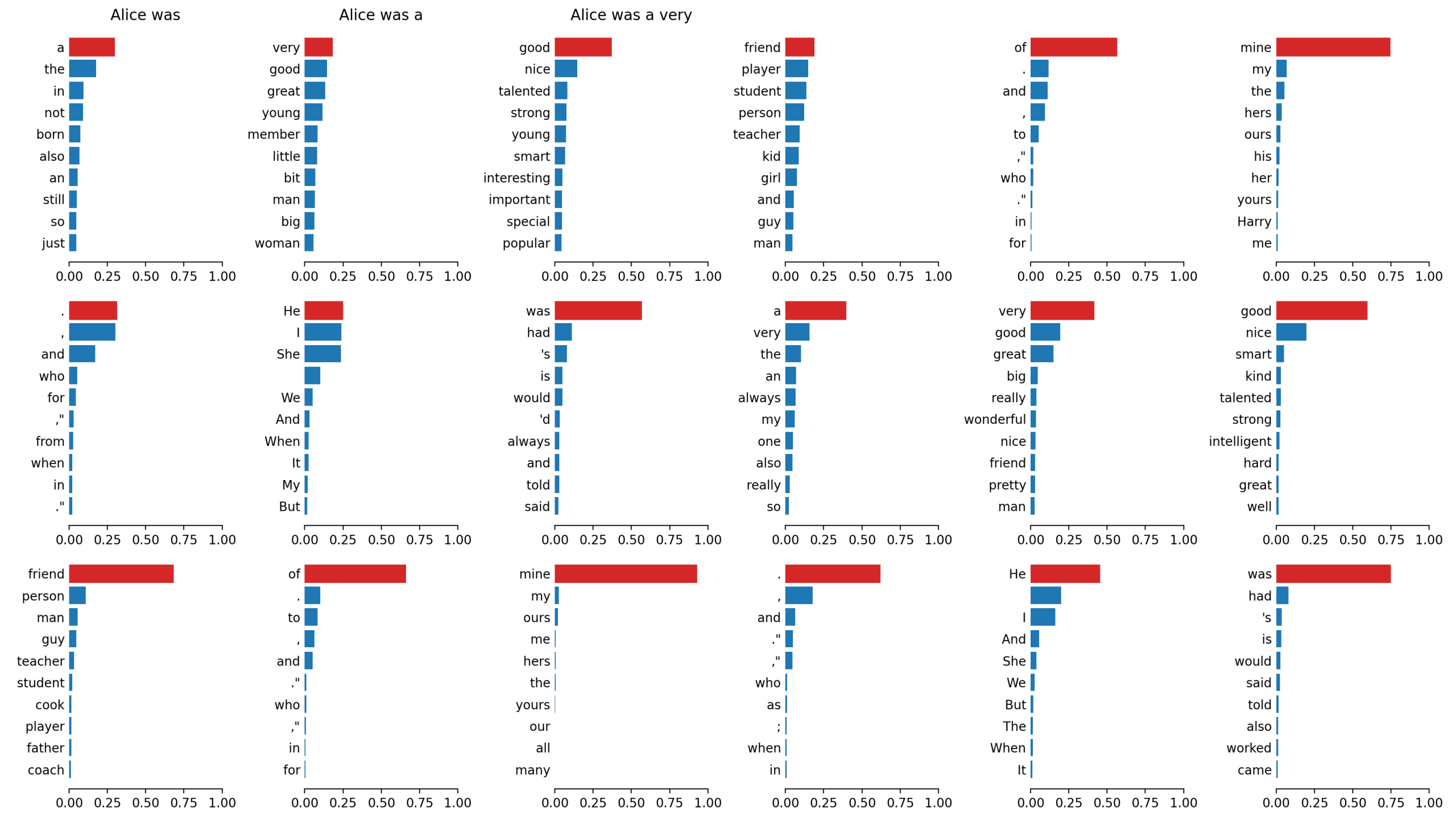


Unused Slides

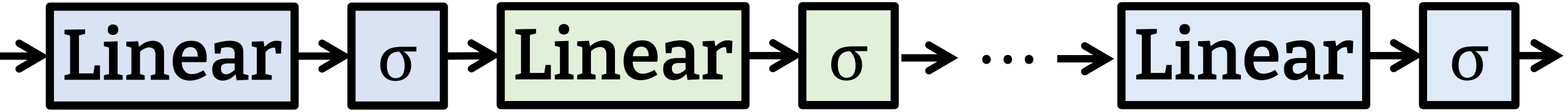

The wizard
's
had
was
ing
is
said
,
of
who
and
The wizard who
had
was
is
has
created
taught
brought
made
killed
gave
The wizard who made
the
it
this
that
him
them
you
his
her
these
spell
first
potion
magic
decision
most
wizard
pact
last
plan
was
,
had
is
would
could
and
made
of
to
make
cure
stop
heal
restore
be
kill
the
save
get
the
her
his
him
them
Harry
himself
Hermione
it
a
world
princess
wizard
girl
boy
day
life
village
man
woman
was
from
,
had
's
is
and
who
.
died
the
her
his
herself
who
was
then
had
save
took
family
children
friends
daughter
baby
child
son
husband
mother
unborn
from
was
,
.
is
had
has
and
?
died
a
the
not
also
named
an
killed
called
one
actually
old
evil
ancient
extremely
expert
elf
eccentric
accomplished
ordinary
excellent
man
wizard
witch
friend
woman
lady
,
-
fellow
and
of
.
,
and
who
from
to
that
named
:
the
mine
his
Harry
hers
Hermione
Dumbledore
her
Lord
King
wizard
family
witch
woman
man
Queen
Lord
Wizard
King
king
0.00
0.25
0.50
0.75
1.00

Most likely next word

Alice was a very good friend of mine. He was a very good friend of mine. He was a very good friend of mine. He was a very good friend of mine. He was a very good friend of mine. He was a very good friend of mine. He was a very good friend of mine. He was a very good friend of mine. He was a very good friend of mine. He was a very good friend of mine. He was a very good friend of mine. He was a very good friend of mine. He was a very good friend of mine. He was a very


Alice was
a
the
in
not
born
also
an
still
so
just
Alice was a
very
good
great
young
member
little
bit
man
big
woman
Alice was a very
good
nice
talented
strong
young
smart
interesting
important
special
popular
friend
player
student
person
teacher
kid
girl
and
guy
man
of
.
and
,
to
,"
who
."
in
for
mine
my
the
hers
ours
his
her
yours
Harry
me
.
,
and
who
for
,"
from
when
in
."
He
I
She
We
And
When
It
My
But
was
had
's
is
would
'd
always
and
told
said
a
very
the
an
always
my
one
also
really
so
very
good
great
big
really
wonderful
nice
friend
pretty
man
good
nice
smart
kind
talented
strong
intelligent
hard
great
well
friend
person
man
guy
teacher
student
cook
player
father
coach
of
.
to
,
and
."
who
,"
in
for
mine
my
ours
me
hers
the
yours
our
all
many
.
,
and
."
,"
who
as
;
when
in
He
I
And
She
We
But
The
When
It
was
had
's
is
would
said
told
also
worked
came
0.00
0.25
0.50
0.75
1.00

MLP: several layers



Contextual embeddings

central bank

river bank

power bank









Token embeddings

Contextual embeddings

It

was

the

best

of

times

Contextual embeddings

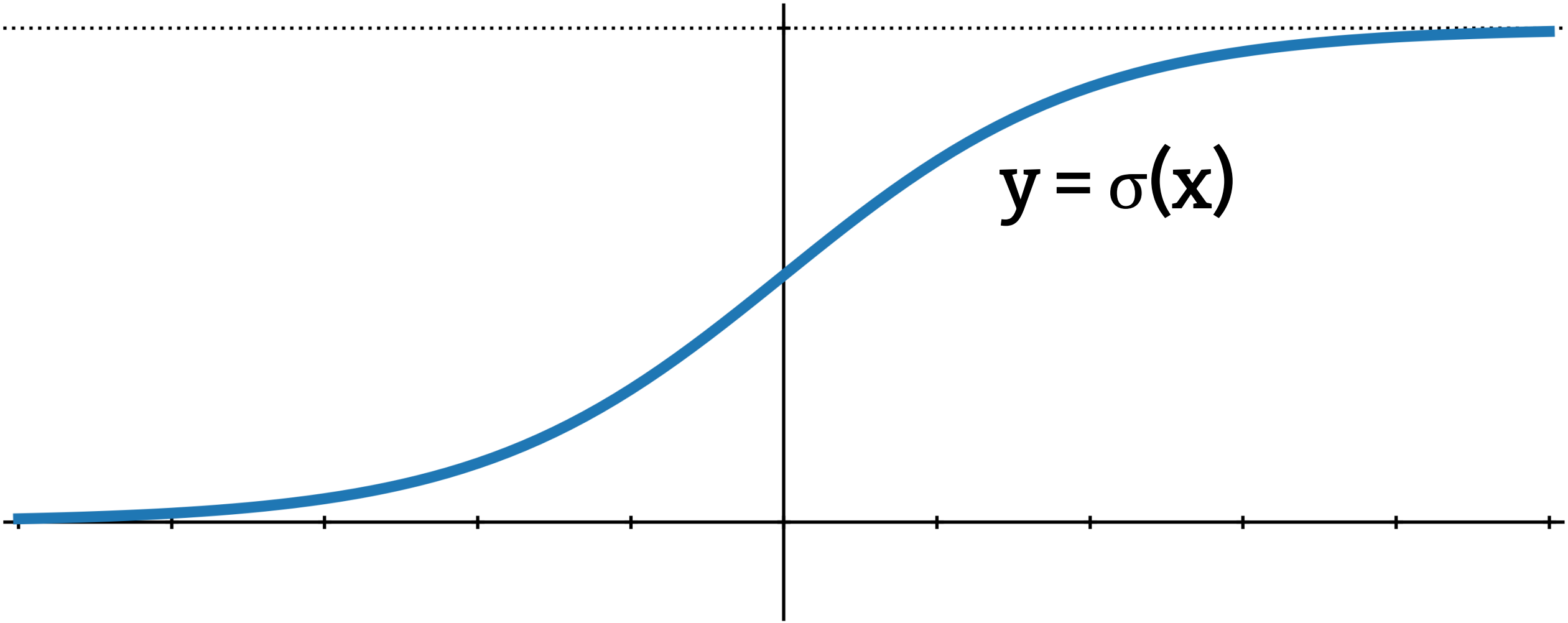
Contextual embeddings

Contextual embeddings

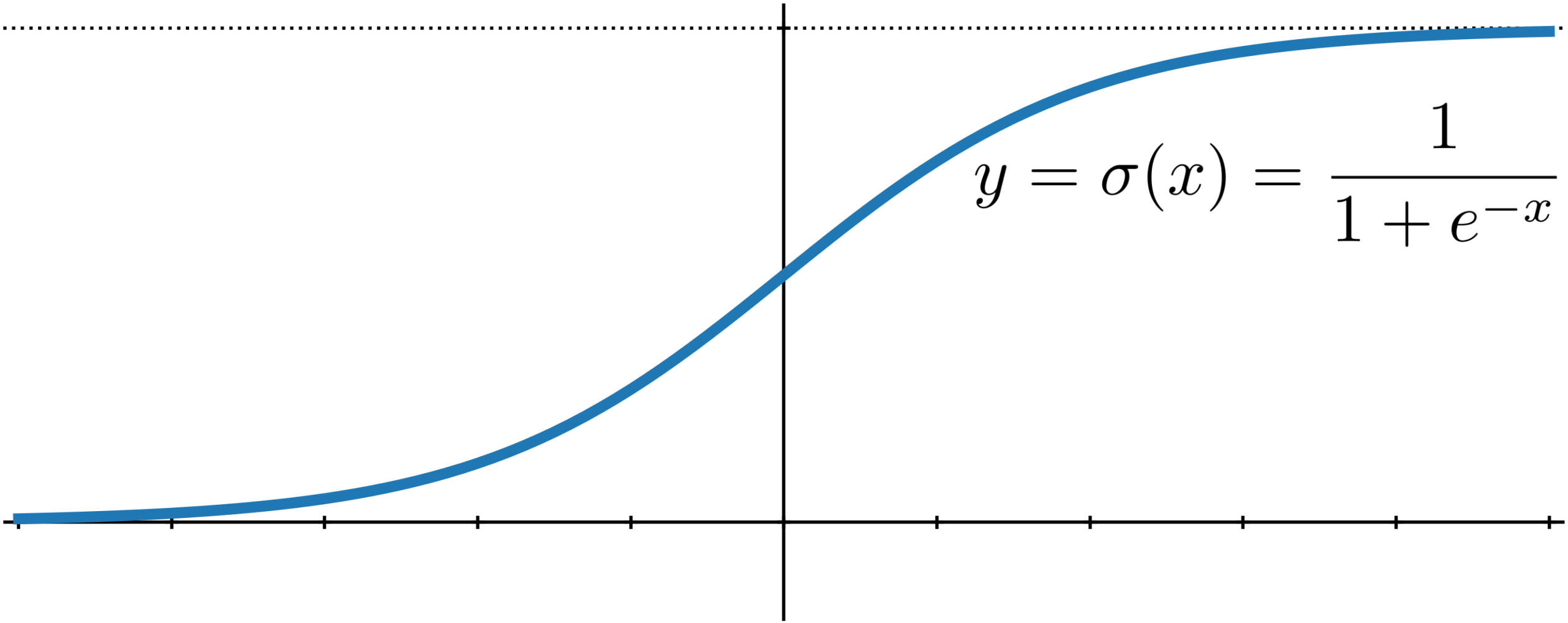
Contextual embeddings

Contextual embeddings

Sigmoid function



Sigmoid function



Training

- TODO?