

AI and Data Analytics

4.3 Self-attention and transformers

Learning language reasoning

Все счастливые семьи похожи друг на друга, каждая несчастливая семья несчастлива по-своему..

Toutes les familles heureuses se ressemblent. Chaque famille malheureuse, au contraire, l'est à sa façon.

Happy families are all alike. Every unhappy family is unhappy in its own way.

Alle glücklichen Familien gleichen einander, jede unglückliche Familie ist auf ihre eigene Weise unglücklich

[L. Tolstoy, 1873]

Examples

If you really want to hear about it, the first thing you'll probably want to know where I was born, and what my lousy childhood was like, and how my parents were occupied and all before they had me, and all that David Copperfield kind of crap, but I don't feel like going into it, if you want to know the truth. In the first place, that stuff bores me, and in the second place, my parents would have two hemorrhages apiece if I told anything pretty personal about them.

VO

Si vous voulez vraiment en entendre parler, la première chose que vous voudrez probablement savoir est où je suis né, et ce que mon enfance moche était, et comment mes parents étaient occupés et tout ce qu'ils avaient avant moi, et tout ce que David Copperfield, mais je n'ai pas envie d'y aller, si tu veux savoir la vérité. En premier lieu, cela m'ennuie, et en second lieu, mes parents auraient deux hémorragies si je leur racontais quelque chose de très personnel.

VF Google

Si vous voulez vraiment que je vous dise, alors sûrement la première chose que vous aller demander c'est où je suis né, et à quoi ça ressemblait, ma saloperie d'enfance et c'est que faisaient mes parents avant de m'avoir, et toutes ces conneries à la David Copperfield, mais j'ai pas envie de raconter ça et tout. Primo, ce genre de trucs ça me rase, et secundo, mes parents ils auraient chacun une attaque ou même deux chacun, si je me mettais à baratiner sur leur compte quelque chose d'un peu personnel.

VF officielle

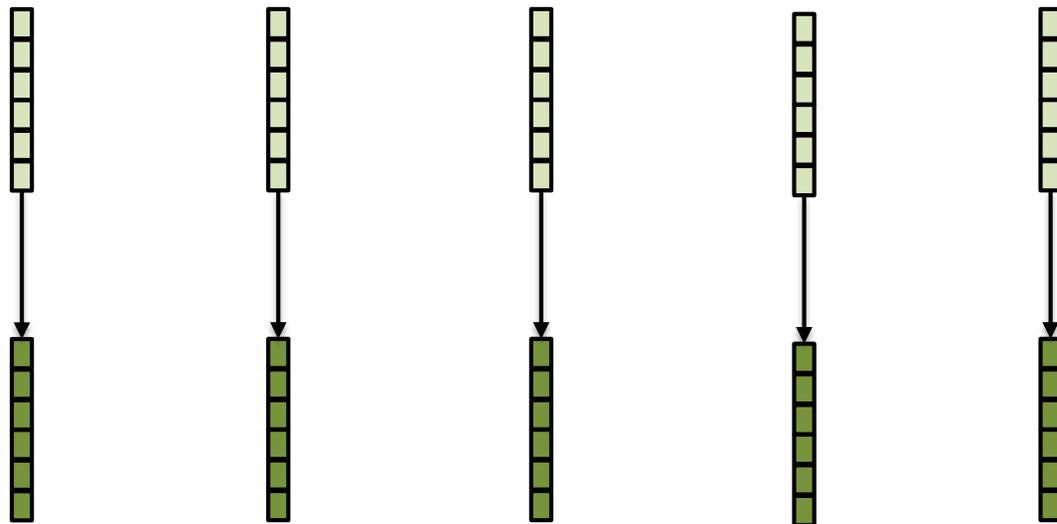
J.D. Salinger, « the catcher in the Rye »

Contextualization

We suppose a set $\mathbf{X} = \{\mathbf{x}_i\}$ of items (vectors).

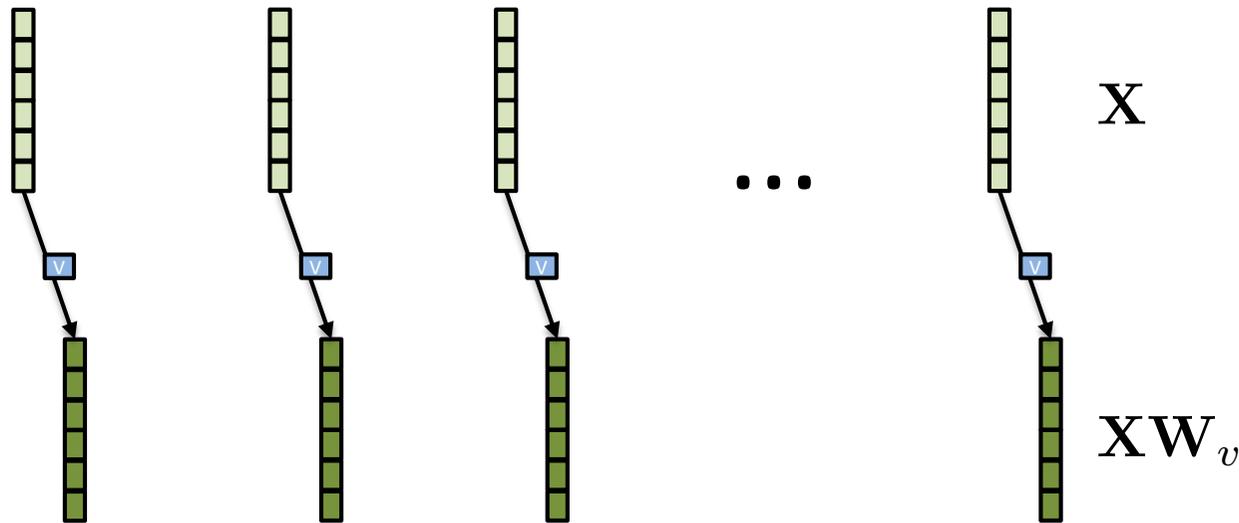
Iteratively "enrich" each item by providing context from the other items

The quick brown fox jumps



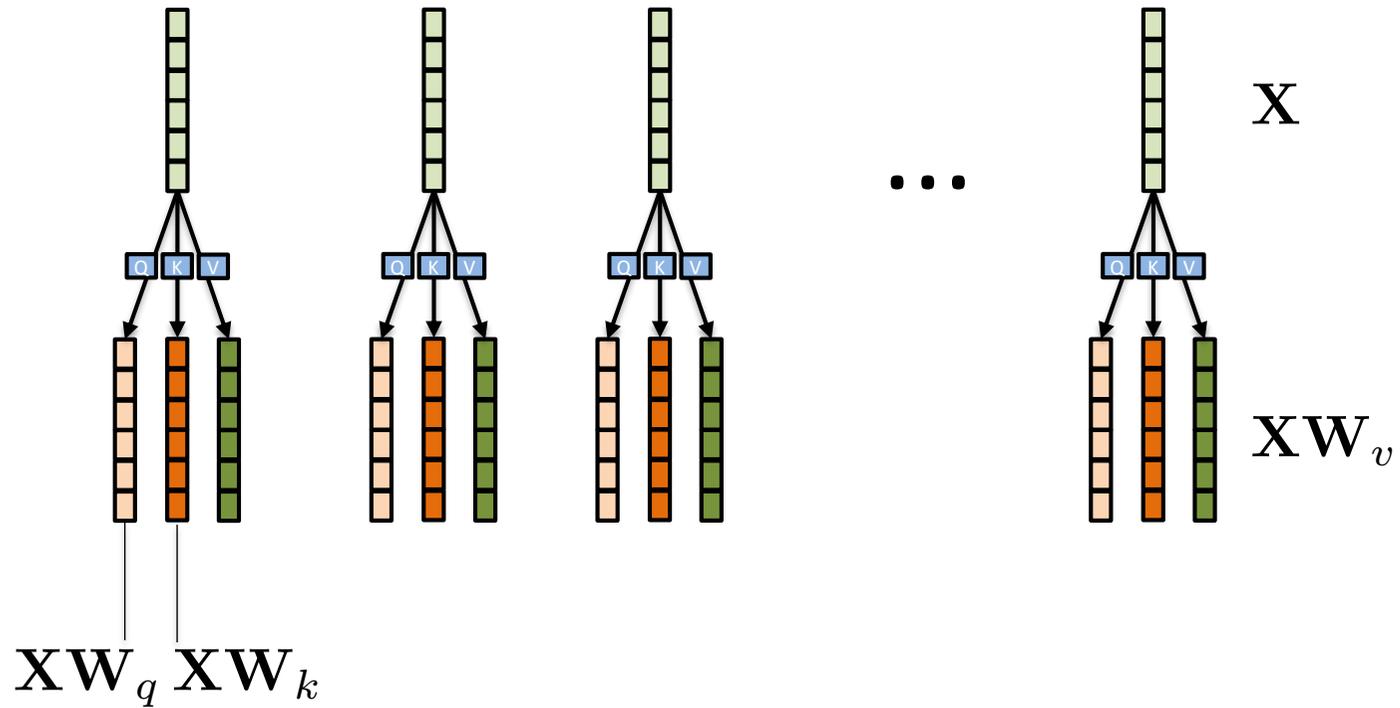
*quick
brown fox*

Transformers: attention is all you need

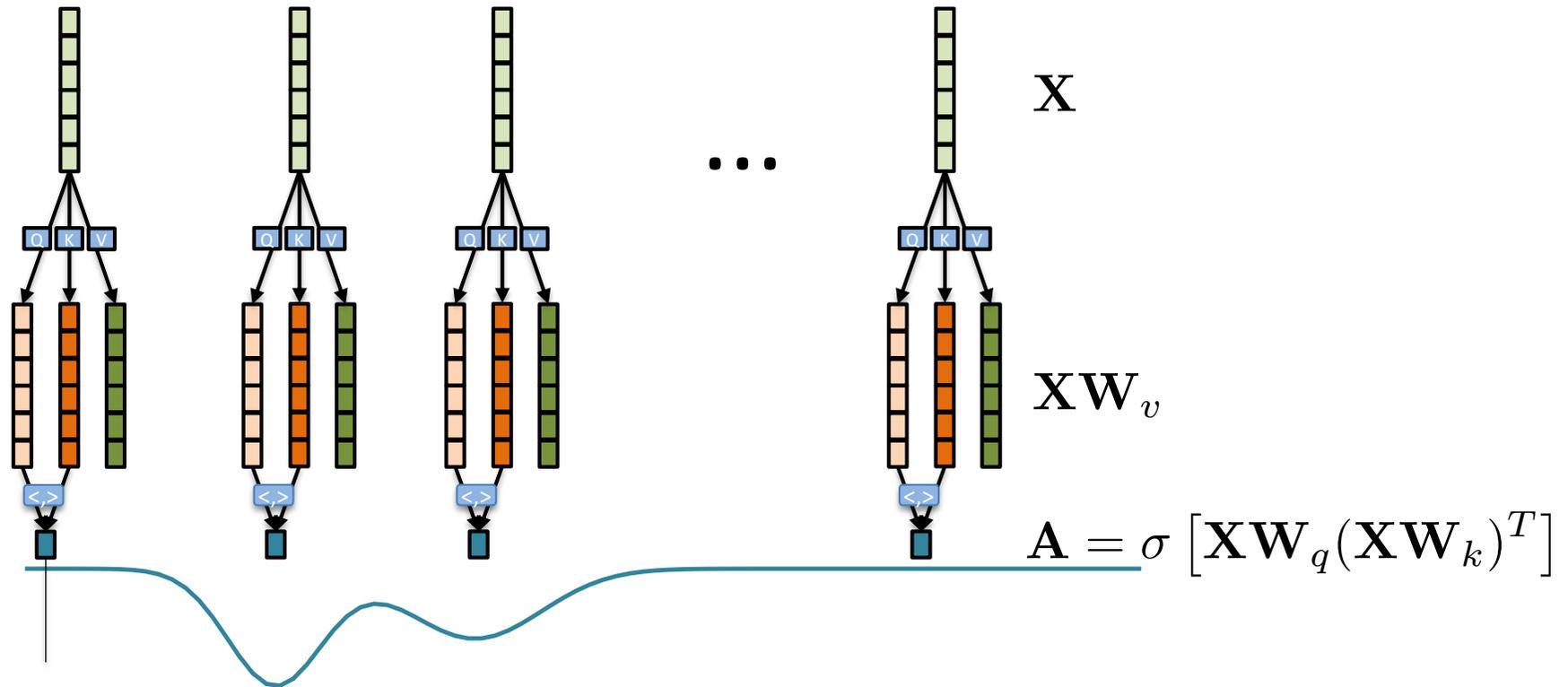


We suppose unordered data (a set) $\mathbf{X} = \{\mathbf{x}_i\}$. Each input item is "transformed" by a linear transform weighted by attention.

Transformers: attention is all you need

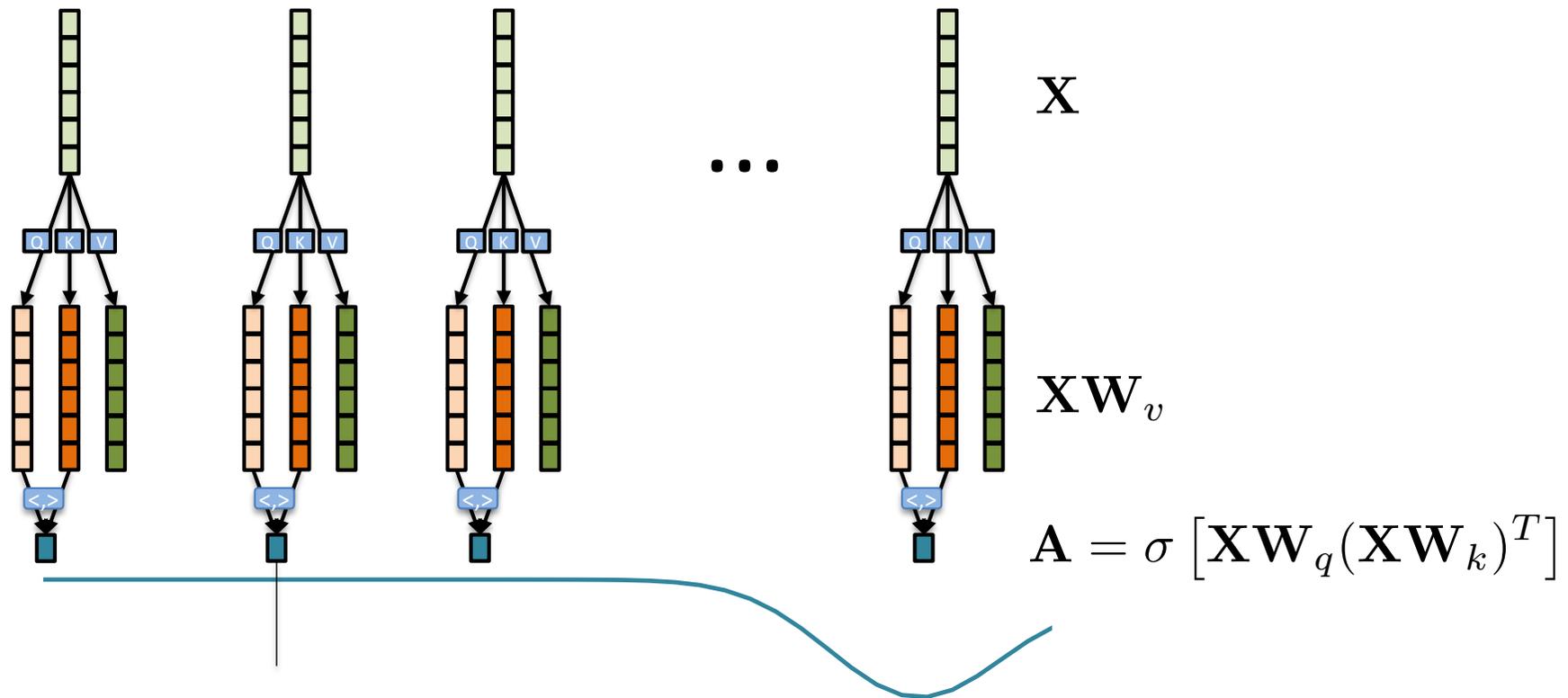


Transformers: attention is all you need



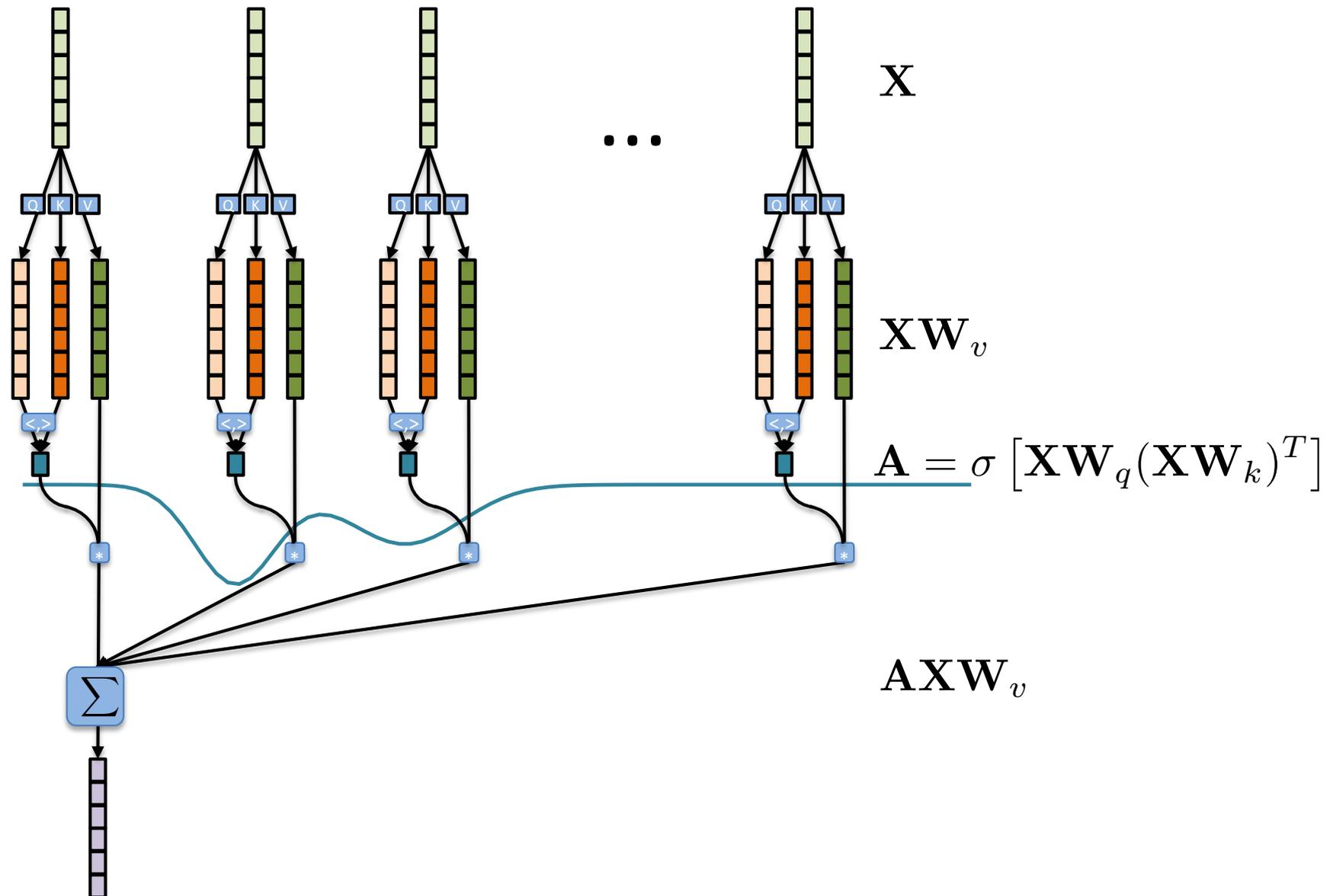
Each item has its own distribution,
associated with its query vector!

Transformers: attention is all you need

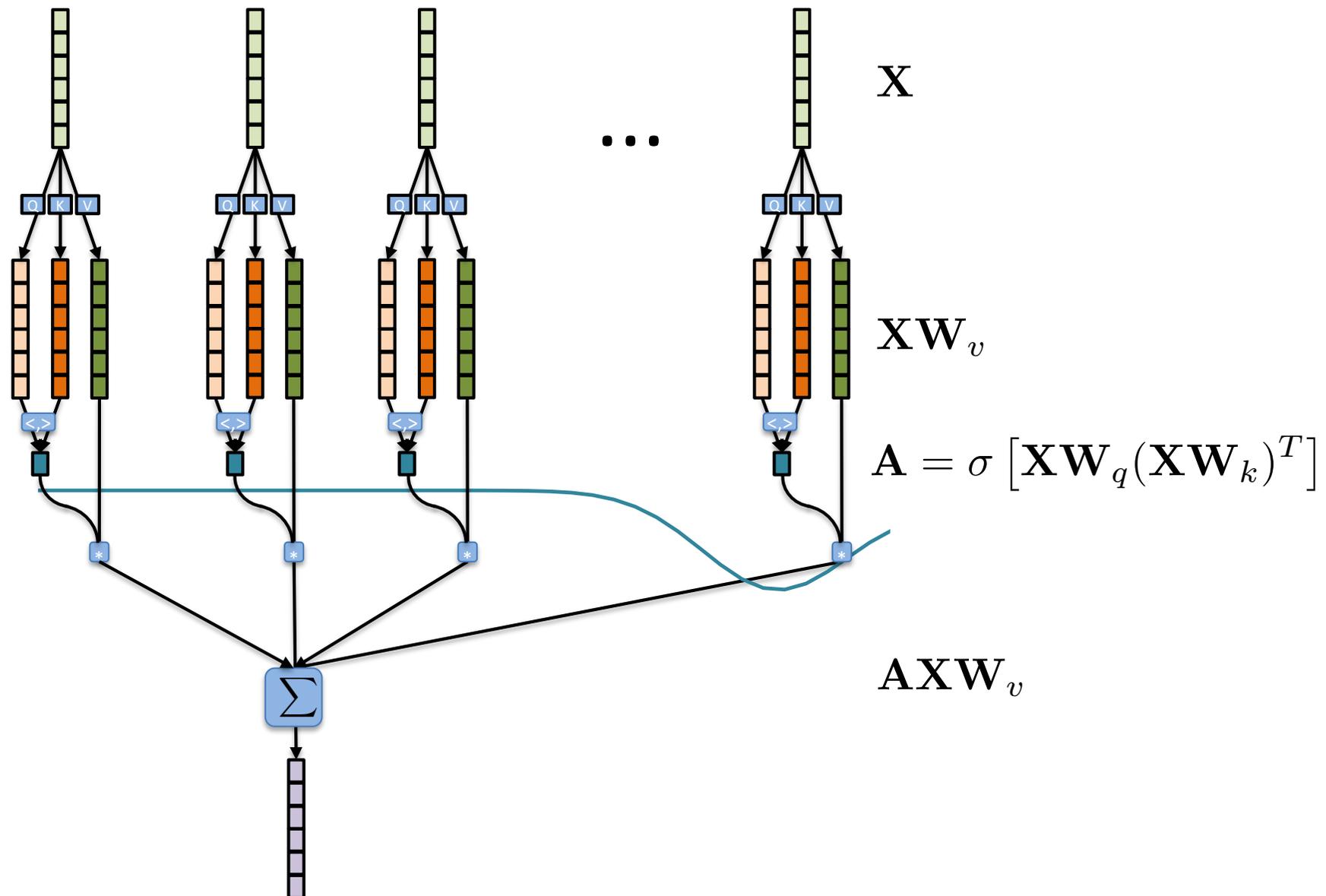


Each item has its own distribution,
associated with its query vector!

Transformers: attention is all you need



Transformers: attention is all you need

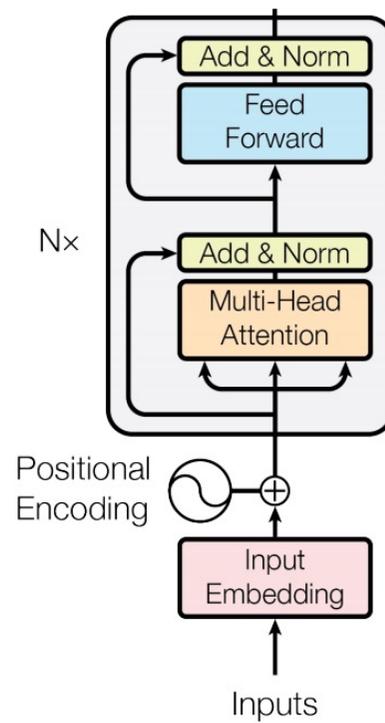


Integration ...

Toutes les familles heureuses se ressemblent. Chaque famille malheureuse, au contraire, l'est à sa façon.



Happy families are all alike. Every unhappy family is unhappy in its own way.



Integration ...

Toutes les familles heureuses se ressemblent. Chaque famille malheureuse, au contraire, l'est à sa façon.



Happy families are all alike. Every unhappy family is unhappy in its own way.

