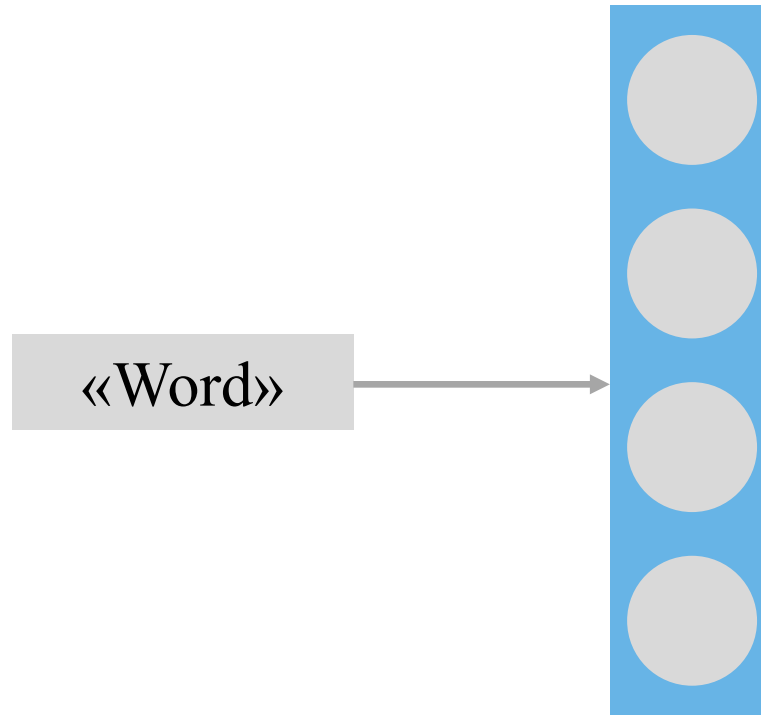
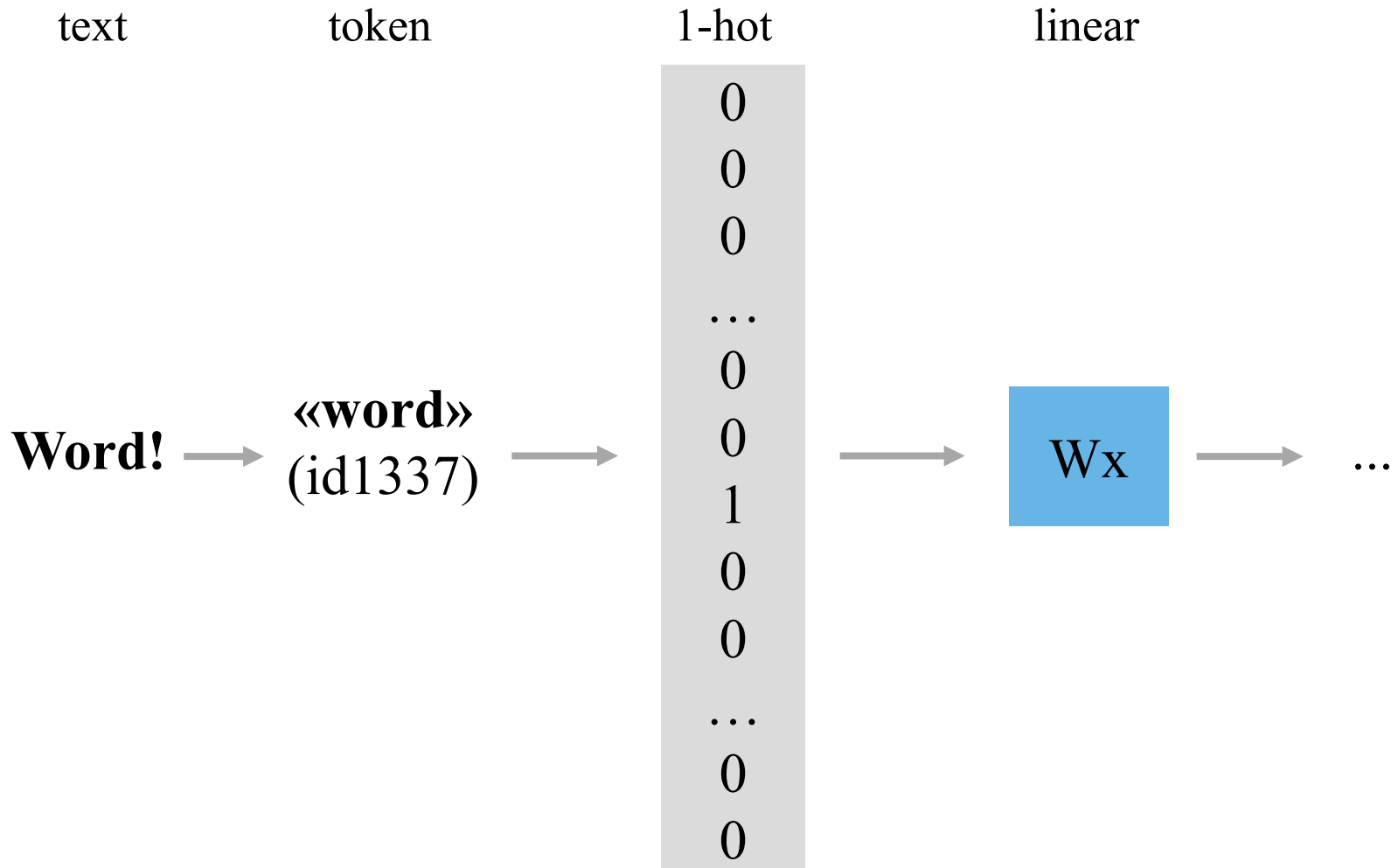


Word embeddings

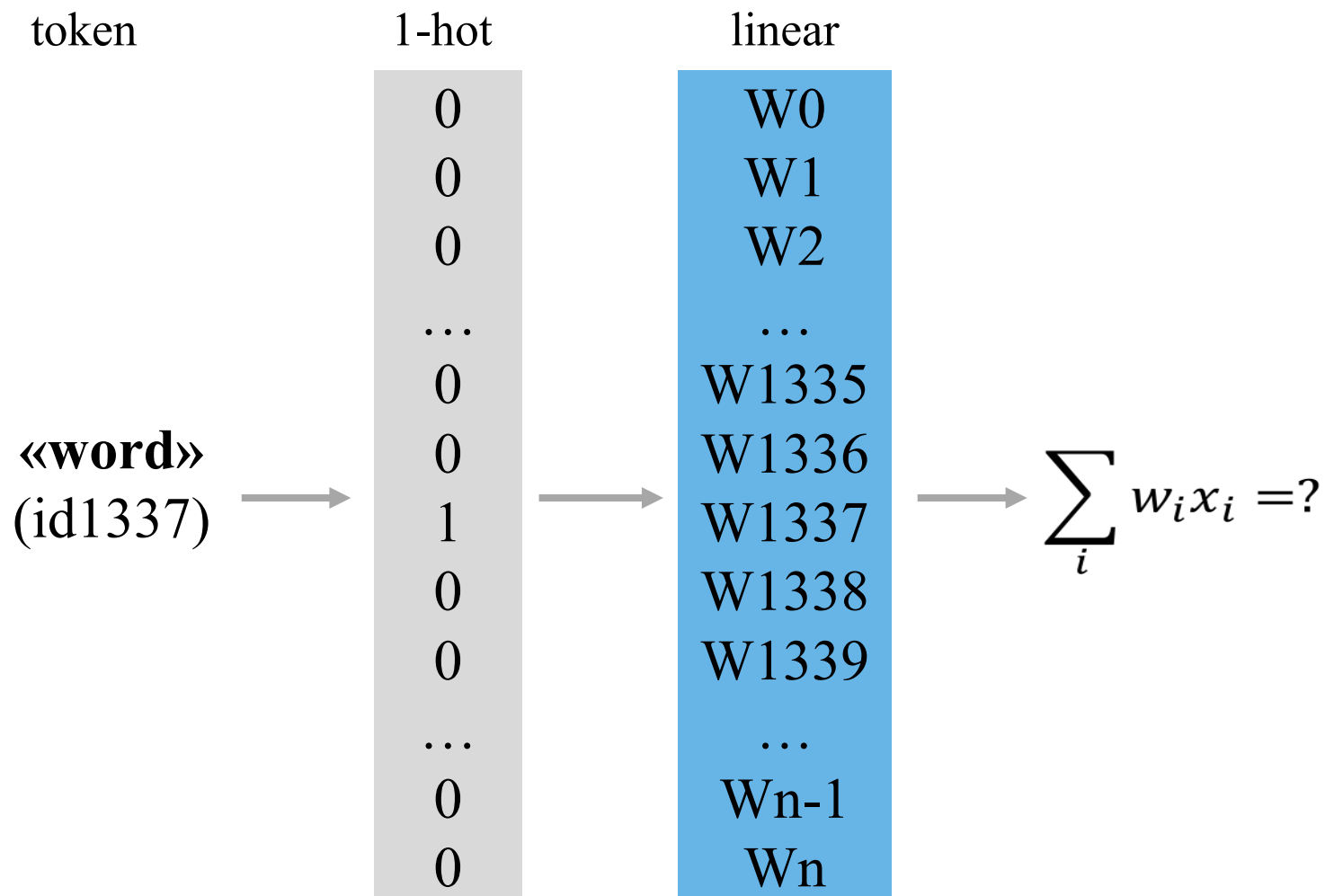
We want a compact representation of text so that we could use it for neural nets!



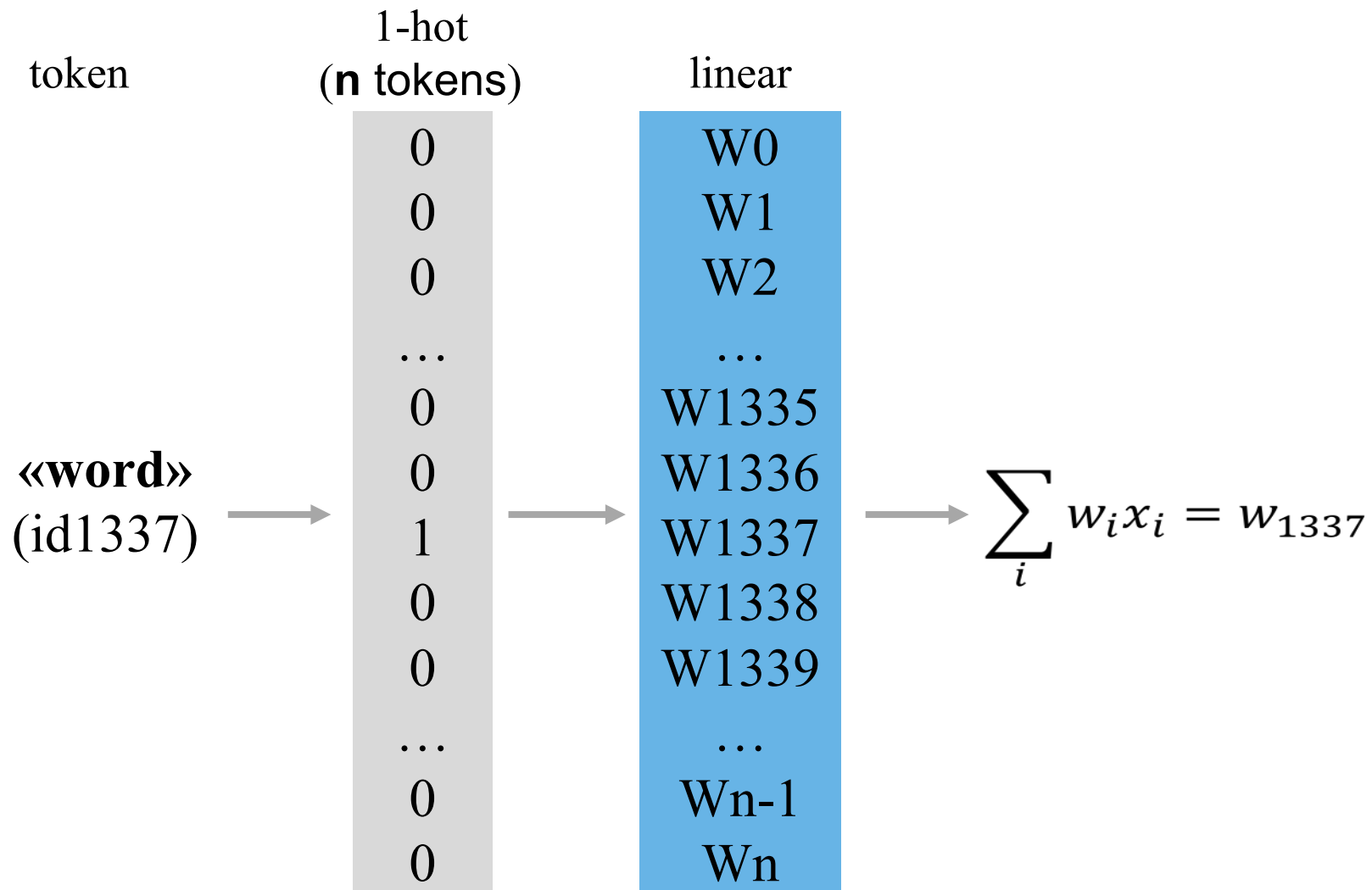
Sparse vector products



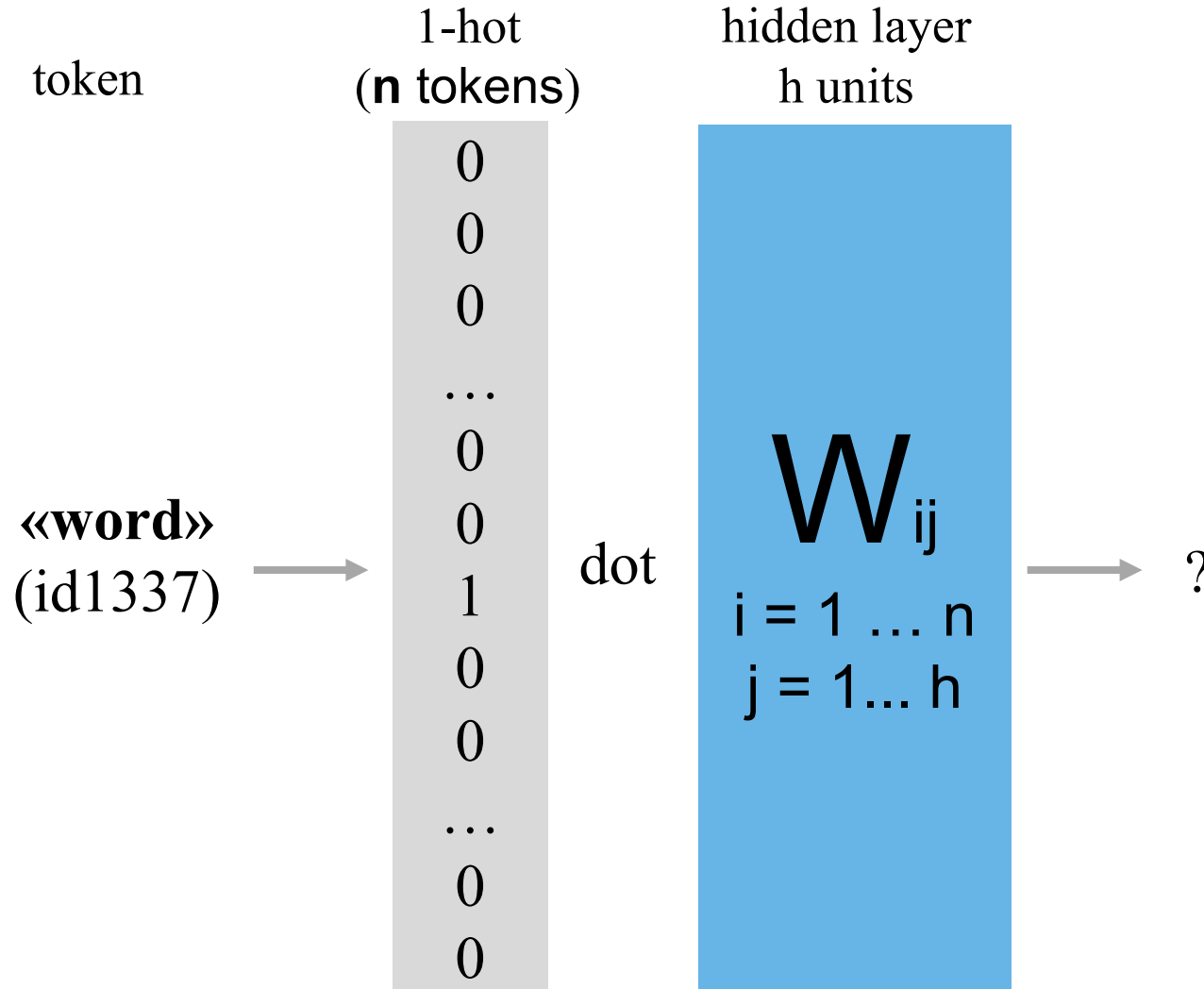
Sparse vector products



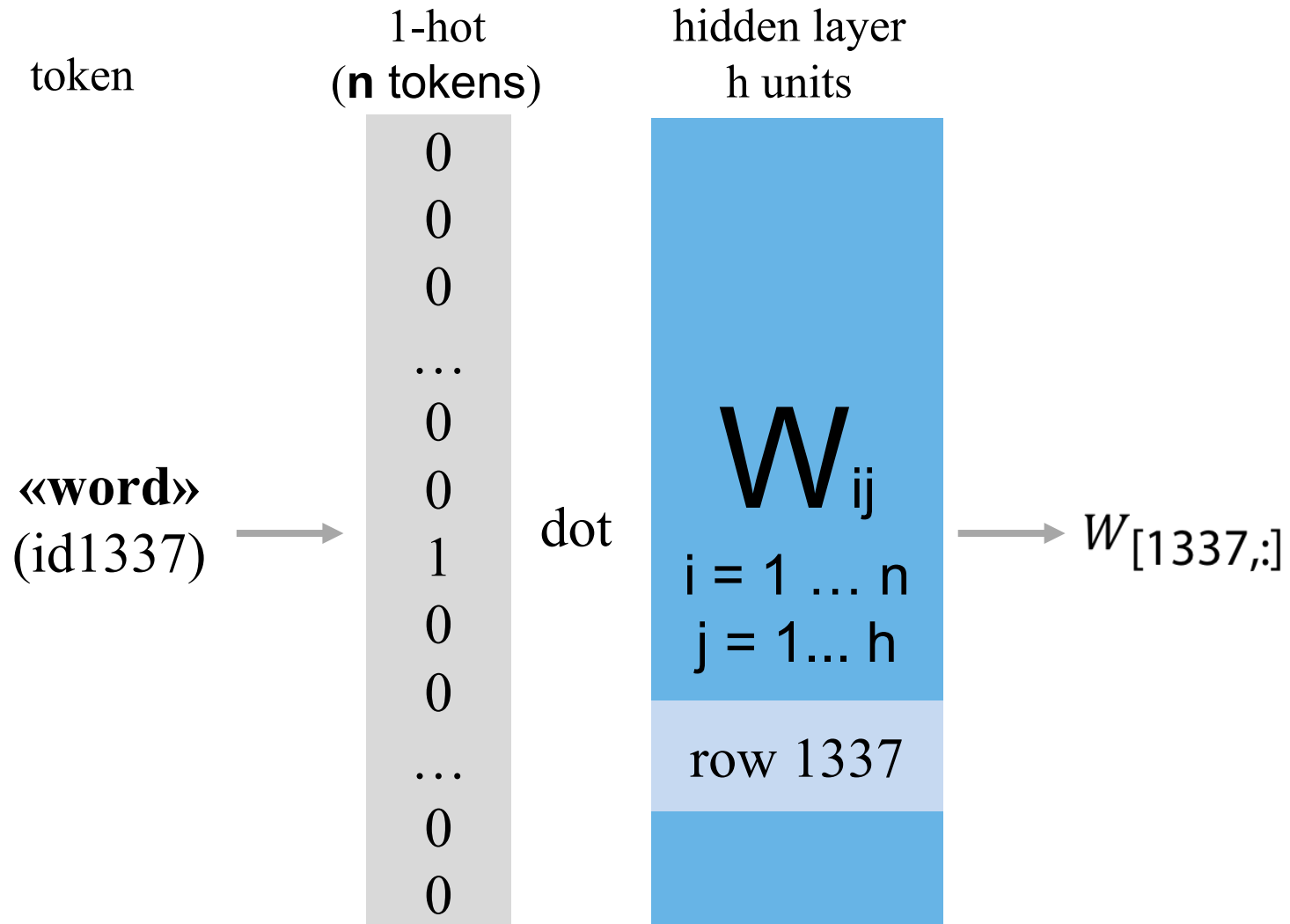
Sparse vector products



Sparse vector products

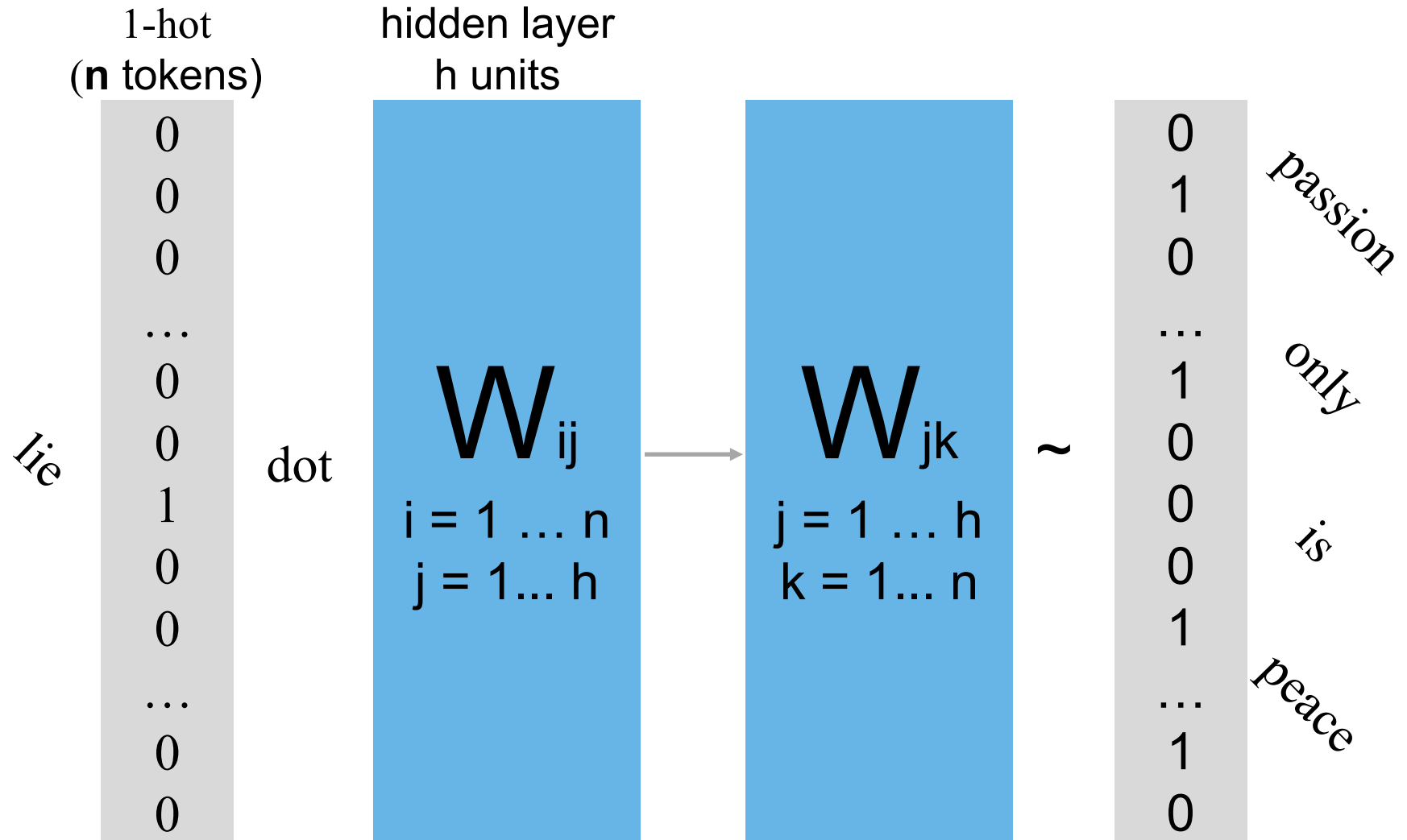


Embedding



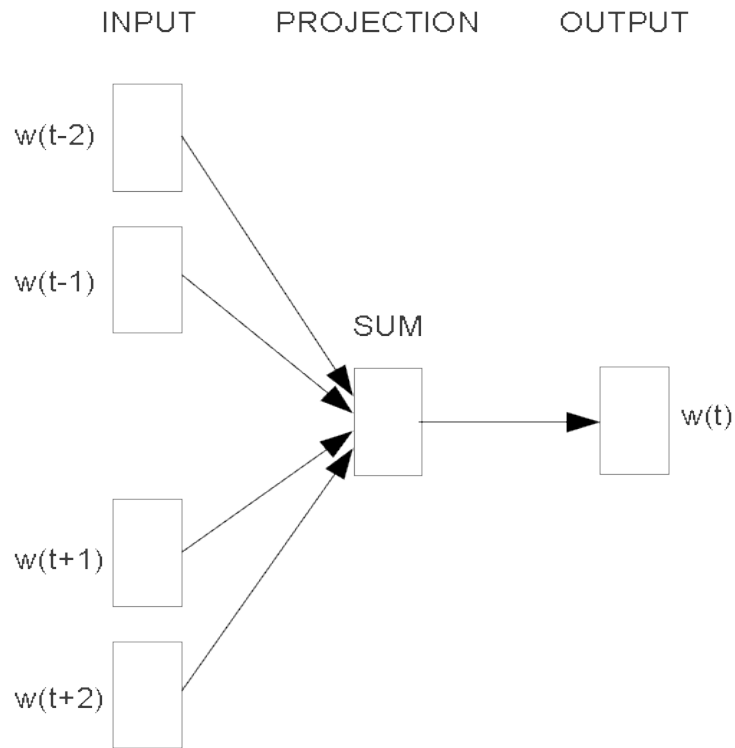
Embedding: word2vec

“Peace is a lie, there is only passion”

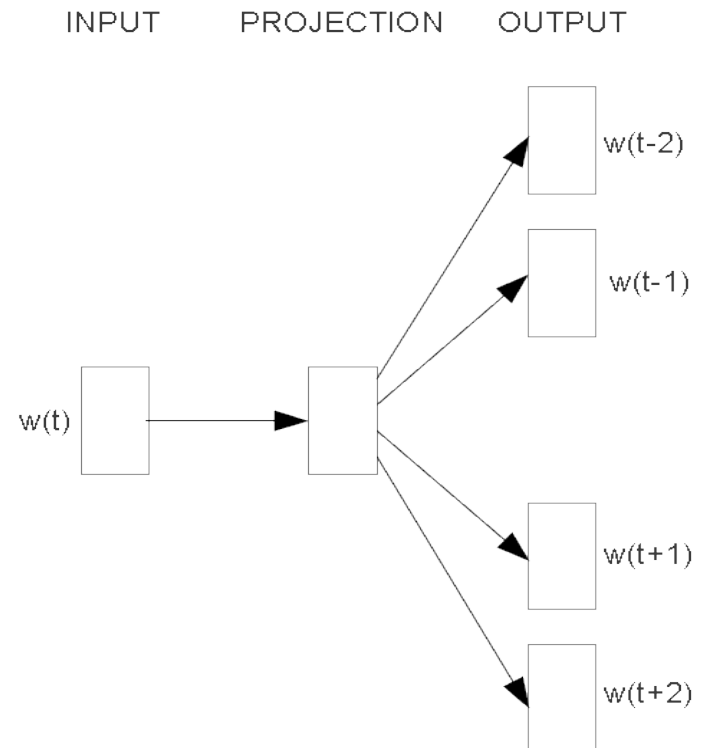


Embedding: word2vec

the *distributional hypothesis* : similar context = similar meaning



CBOW



Skip-gram

Embedding: word2vec

Side effect: synonyms

“nice” ~ “beautiful”

“hard” ~ “difficult”

Embedding: word2vec

Side effect: synonyms

“nice” \sim “beautiful”

“hard” \sim “difficult”

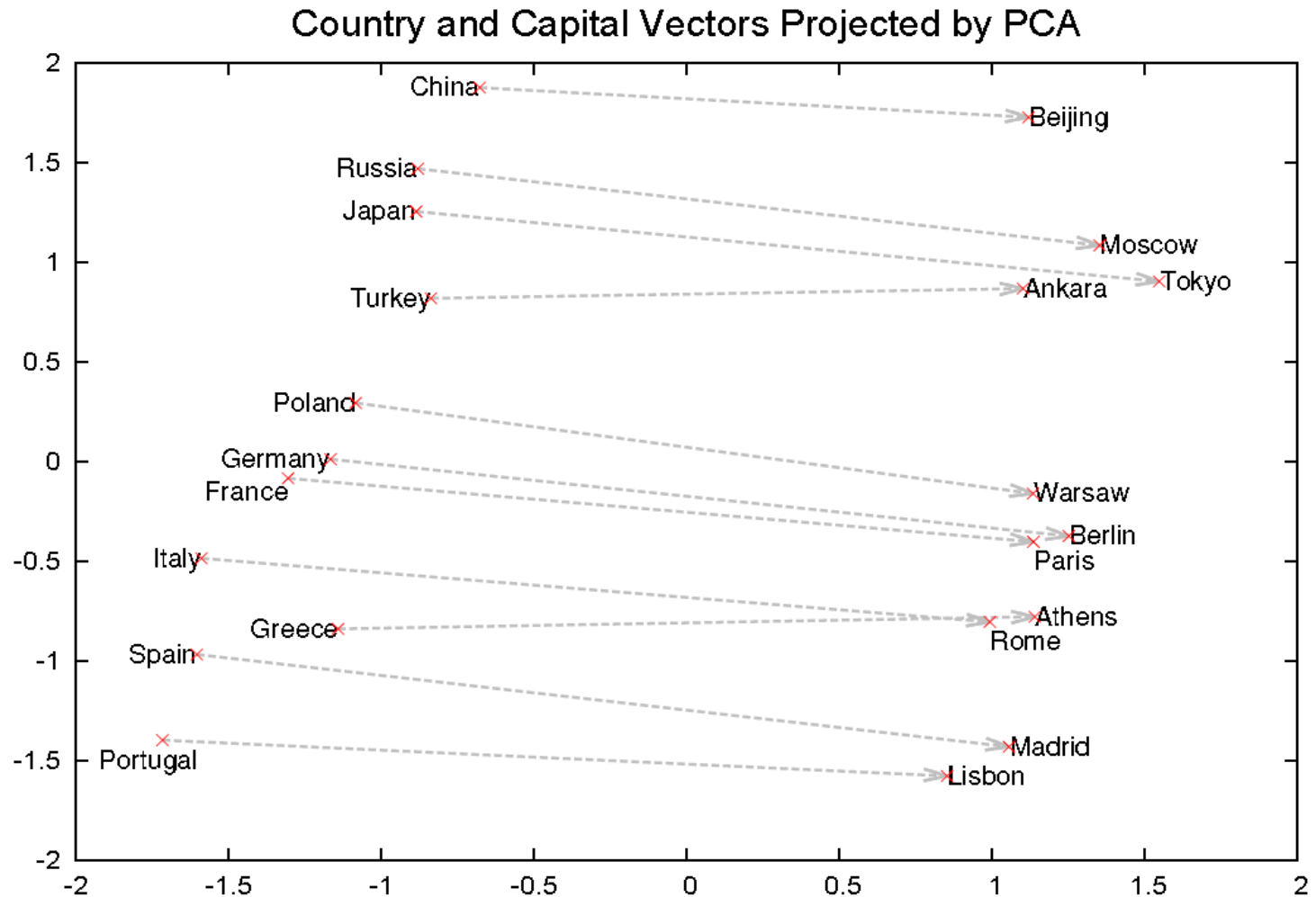
Side effect: word algebra

“king” - “man” + “woman” \sim “queen”

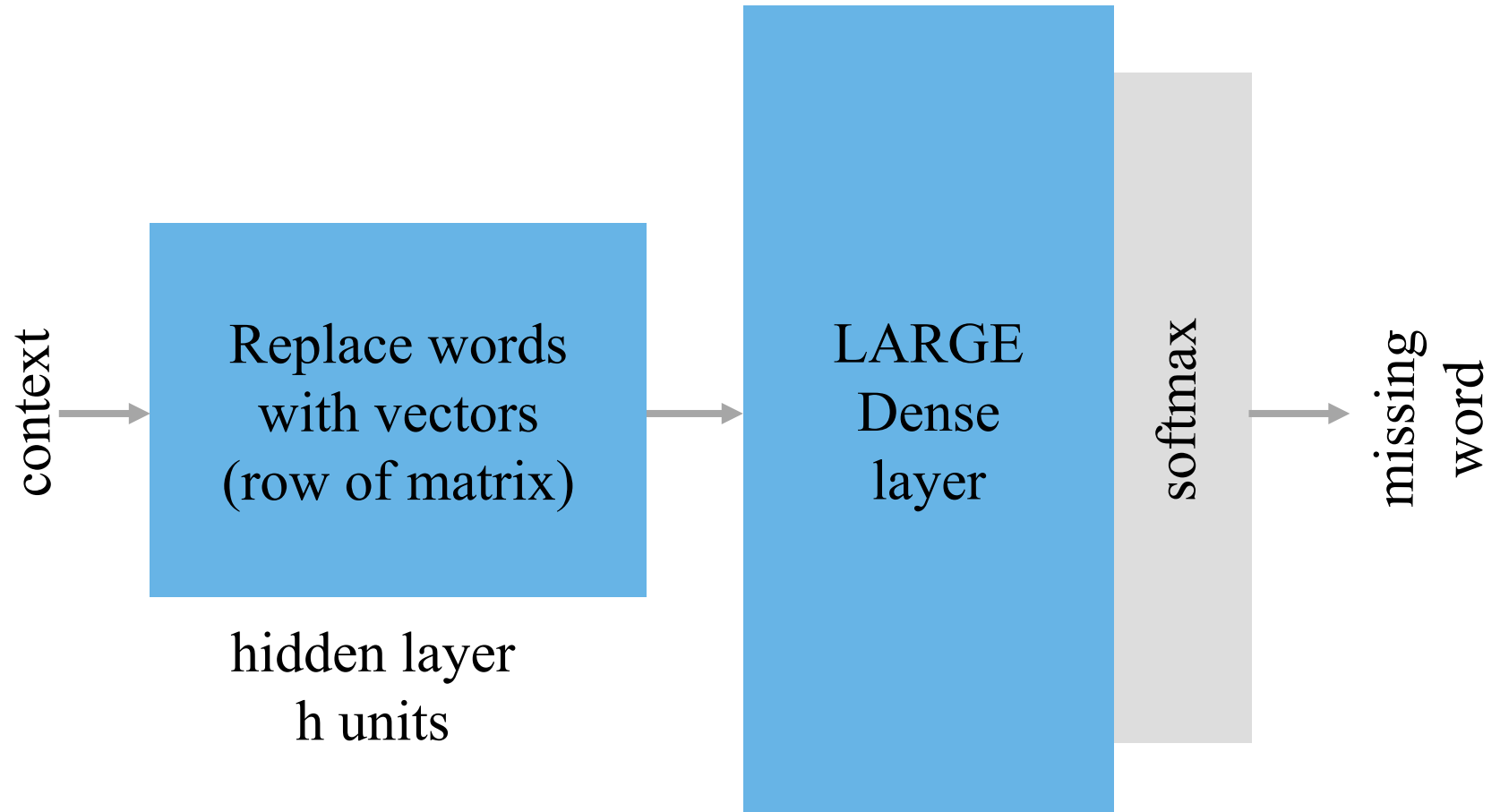
“moscow” - “russia” + “france” \sim “paris”

Embedding: word2vec

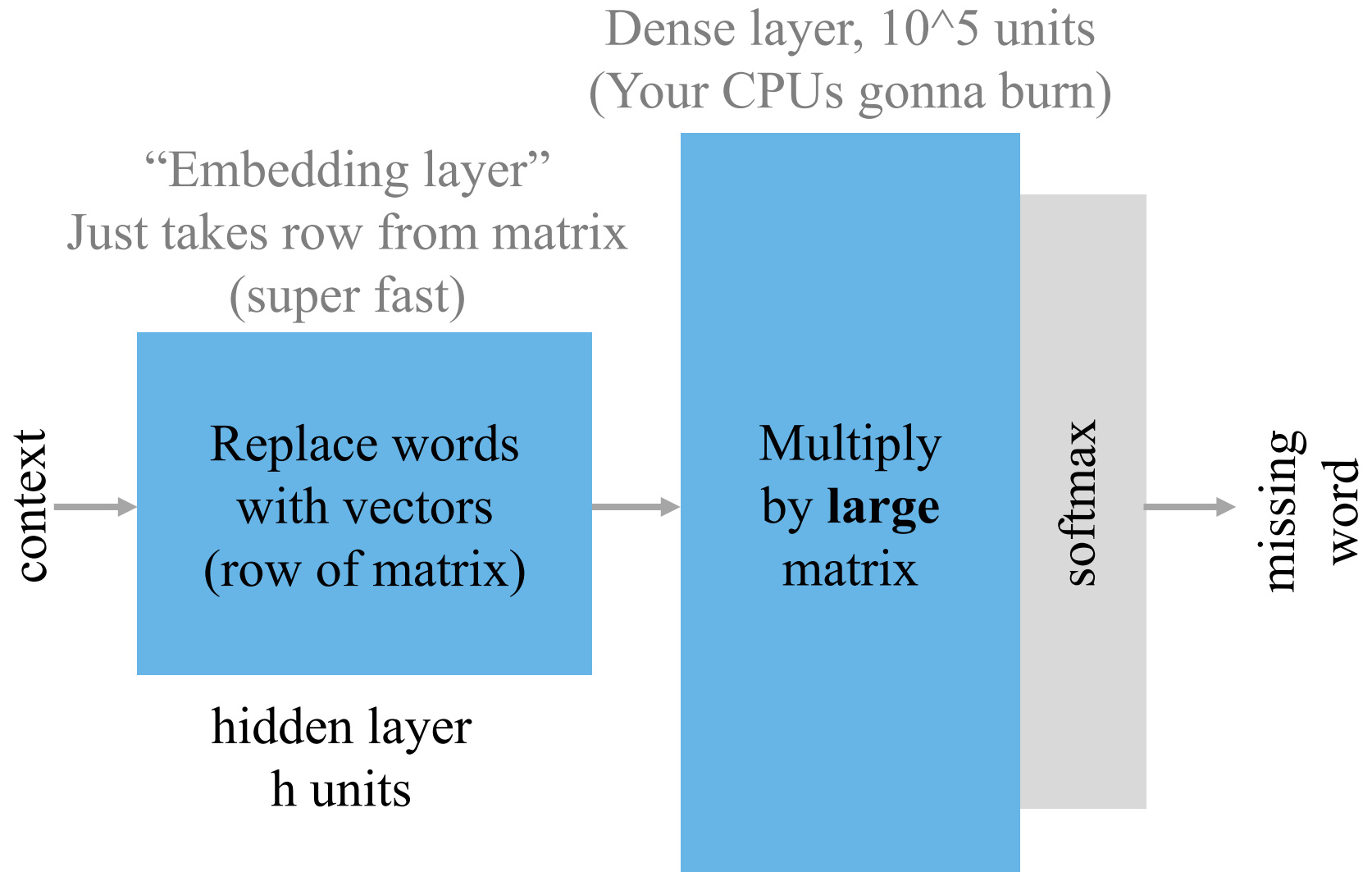
Side effect: word algebra



Softmax problem



Softmax problem



More word embeddings

Faster softmax:

- Hierarchical softmax, negative samples, ...
- learn more

More word embeddings

Faster softmax:

- Hierarchical softmax, negative samples, ...
- learn more

Alternative models: GloVe

More word embeddings

Faster softmax:

- Hierarchical softmax, negative samples, ...
- learn more

Alternative models: GloVe

Sentence level:

- Doc2vec, skip-thought (using rnn)

More word embeddings

Faster softmax:

- Hierarchical softmax, negative samples, ...
- learn more

Alternative models: GloVe

Sentence level:

- Doc2vec, skip-thought (using rnn)

To be continued...
in the NLP course