

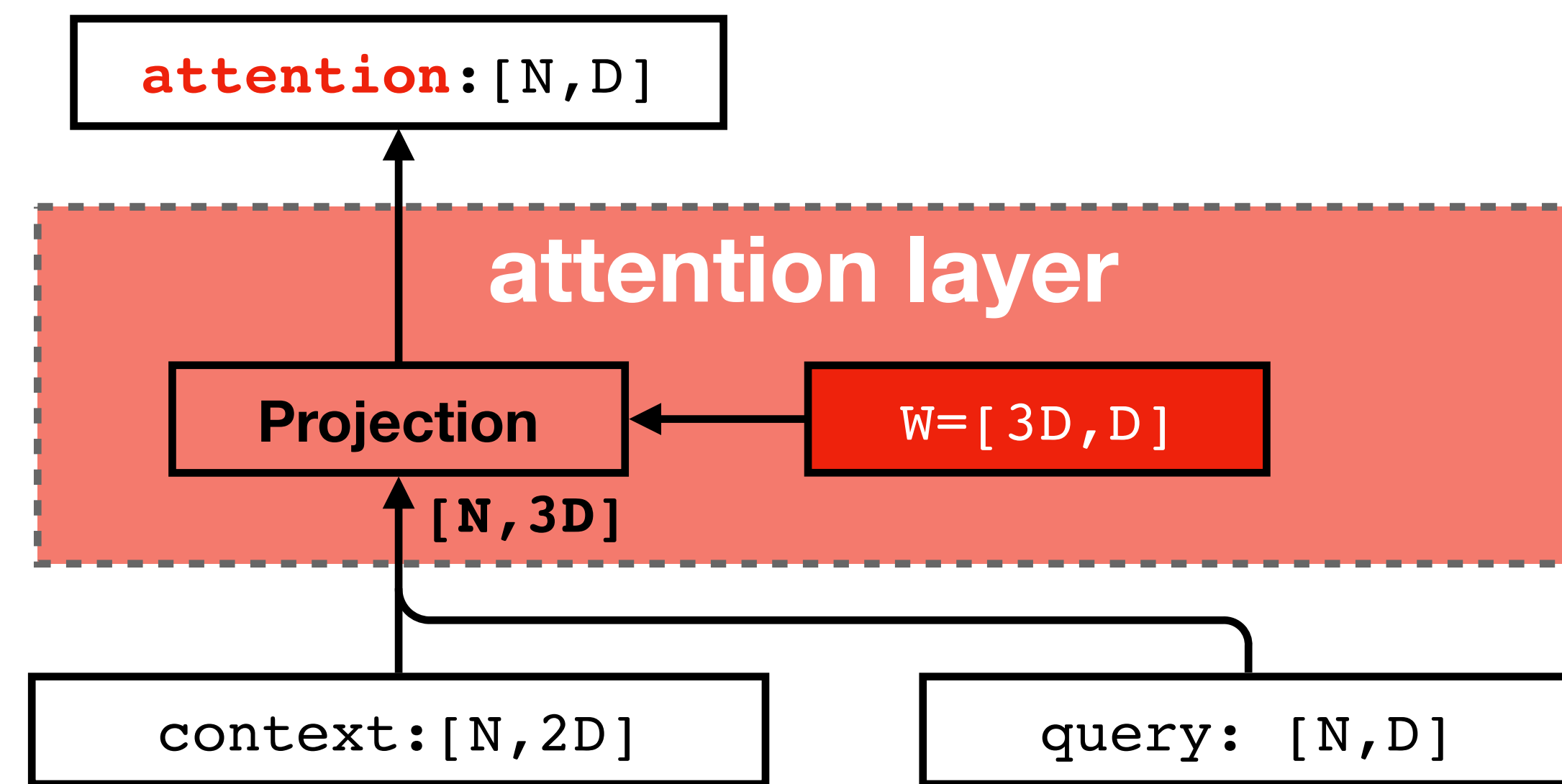
```

keys      = encoder_outputs.reshape([Tsrc*N, 2*D]).dot(W).reshape([Tsrc, N, D])

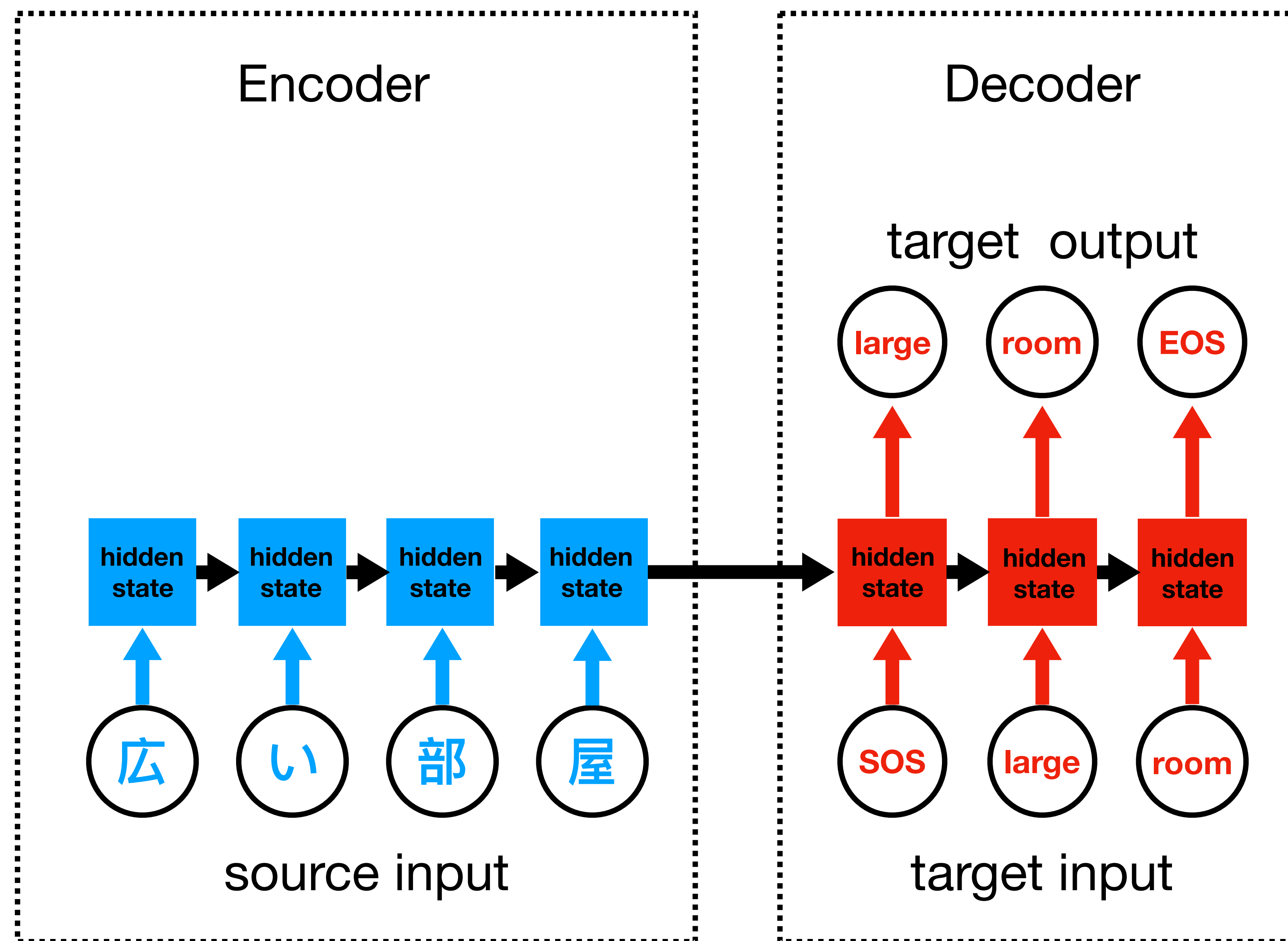
raw_scores = (keys*query).sum(axis=2) # dot product

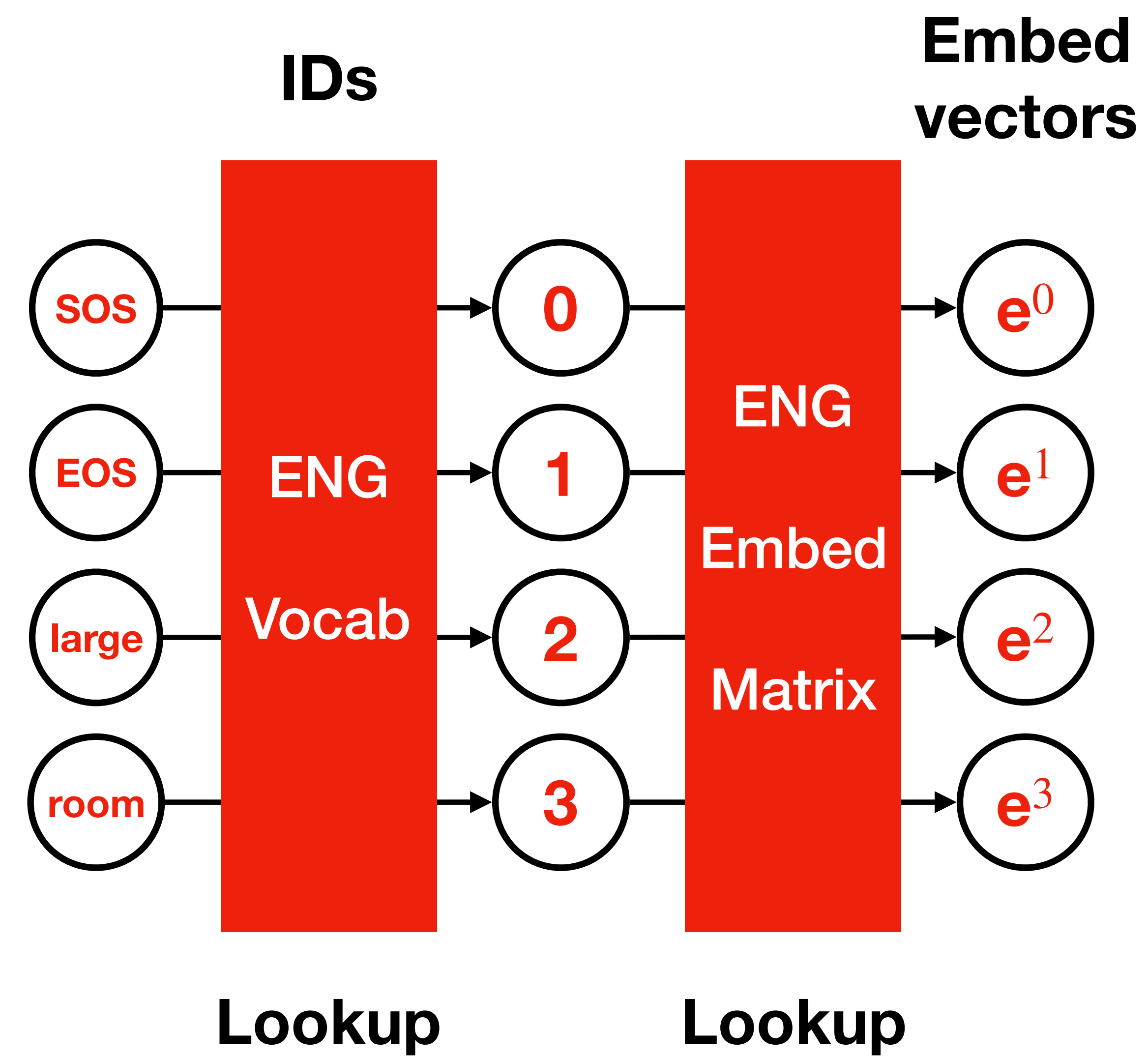
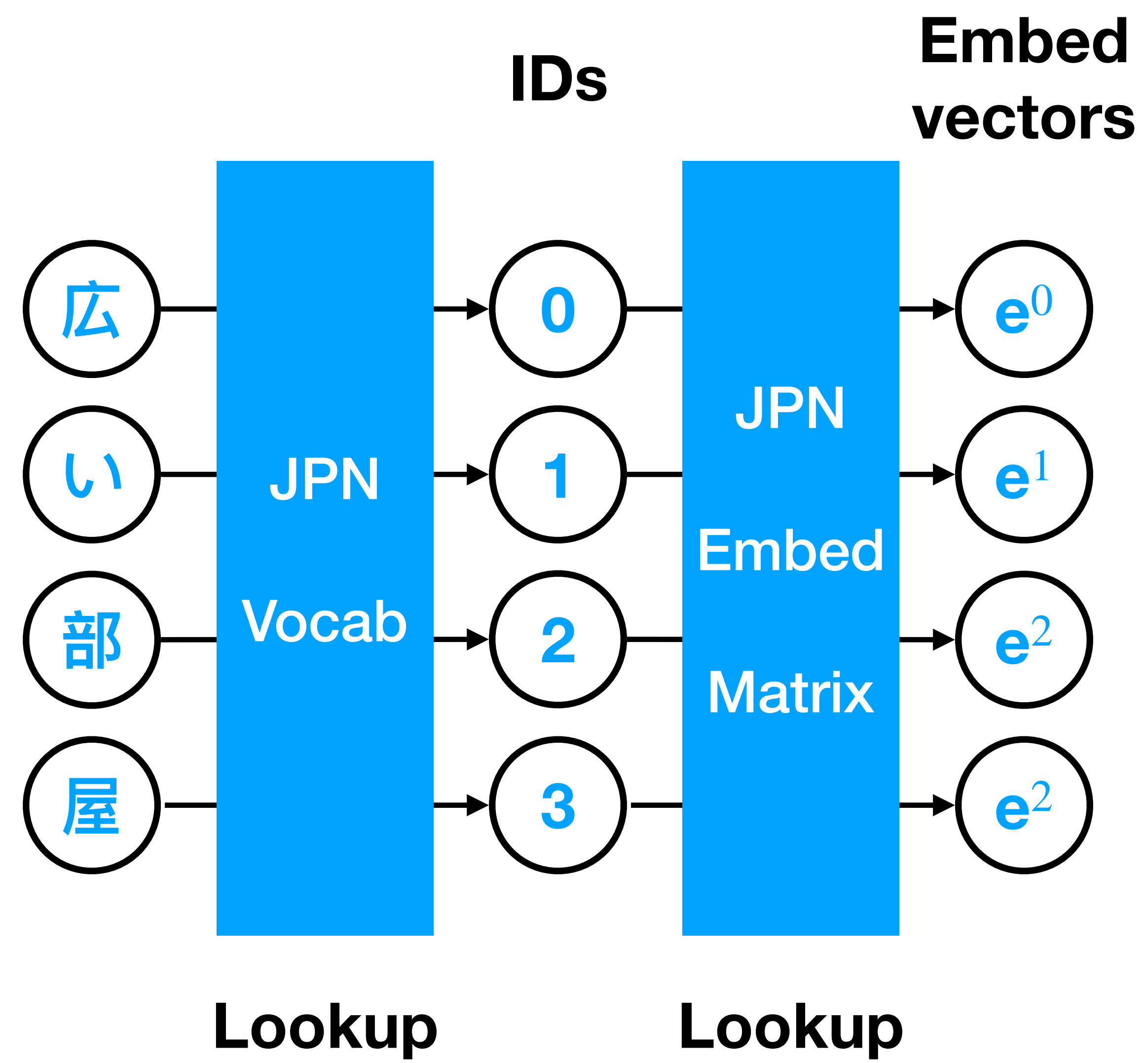
scores    = np.exp(-raw_scores) / np.exp(-raw_scores).sum(axis=0) # softmax

context   = (encoder_outputs * np.expand_dims(scores, axis=2)).sum(axis=0)
  
```



```
attention = np.hstack([context, query]).dot(W)
```

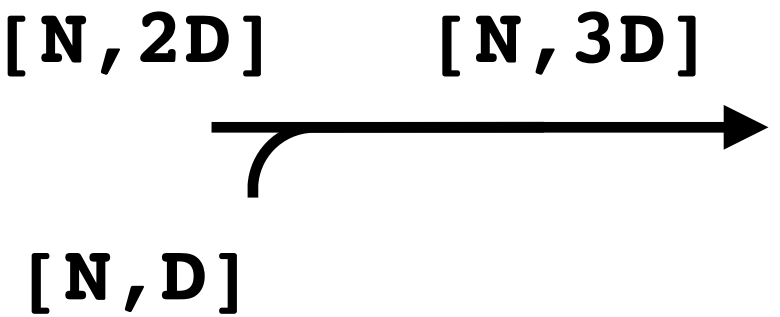




Zero initializer



Concatenate



Share

