2.a

We realize the equivalence that , then we have:

Then we have the linear format that we want, where:

(In the proof, we used the property: without proof, but it could be easily shown by expanding the right hand side of the equality. )

2b

Again, we realize the equivalence that:

.

We also notice that we could change the product, , to the product of a different base, , where denotes the whole vocabulary( is the word in the vocabulary), and denotes the number of occurrences of in . We see that:

This is because for any word in , we multiply to the total result for each occurrence. This is equivalent to multiplying to the power of the total occurrence to the result.

Then, we would denote the representation of input as , where denote the number of occurrence of word in the input . The above equation becomes:

Continue our equivalence relation, we have:

We again realize that the first term is simply the expanded version of a dot product. We could express this as a linear combination where we define as , and .