**Answer to the question 2**

From response to the question 3, we have found that for (hats, fedoras) and (cats, felines) the probabilities are increasing as expected.

However, for between (kids, children), (like, adore), (like, children), (kids, adore) the probabilities are also increasing by same amount though the latter two pairs are wrong translation. Which means for these 4 tuples the translation model is confused.

Based on the above observations, following will be the converged translation probability:

Table

Description automatically generated

From the above table we can see that the translation model failed here. It can translate (hats, fedoras) and (cats, felines) but it cannot translate the other words.

The model is confused between (kids, children) and (like, adore). This is because in the dataset we have only one sentence of these two pairs. Consequently, the translation model cannot distinguish between them. Based on this observation, if a dataset does not have multiple sentences for same pair of word it will fail in the way it fails here.