Let us call the matrix for T1 and T2 as M1 and M2.

T1 = abcd

T2= acdea

Union (T1,T2) =abcde

In T1 , ab =1 ,bc=1,cd=1 (talking 2 alphabets as a pair)

In T2, ac=1,cd=1,de=1,ea=1

Now constructing a labeled complete graph for T1 based on the set of nodes ={a,b,c,d,e}

For simplicity I am showing the edges from “a” only and initializing with 1 on the edges. Then we add the frequency of occurrences ..in this case only “ab” occurs 1 so it becomes 2. Then we divide with the sum.

Which is 2+1+1+1=5. Now do the same process for each of the other nodes. This will be the Markov chain for T1…Call it M1 Similarly you will have a Markov chain for T2..call it M2. Then you will compute KLD based on the matrix values of M1 and M2.

Example https://www.statology.org/kl-divergence-python/