Assignment 2 - CS753

Ankur Verma (20D070014)
Chinmay Pimpalkhare (200100115)
Saksham Katiyar (20D100022)
Vikhyat Agrawal (200260058)

April 2024

Part II

Part C

Constrained beam search: The task was to modify the beam search decoding routine such that the output prediction is guaranteed to have **at least one English token.**

Solution: We solve this problem by proposing a slight modification to the beam search algorithm, such that along with the most probable candidates, we also force a candidate which satisfies the given constraint (regardless of the likelihood of the forced candidate). This ensures that there is always one candidate present in the beam size which satisfies the given constraints. At the last time step, when choosing between the finished sequences, we then pick the best possible sequence, whether or not that sequence has an English alphabet. This is done because we observed that not all utterances have an English letter in their transcription, and that we would be forcing an unlikely candidate to be picked at the time of inference. This is why we choose the most likely sequence at the last time step. A step-by-step guide has been given below to illustrate our solution.

As illustrated in 1, for each sequence candidate, we choose (Beam Size + 1) possible candidates, with the extra 1 being from the constraint which we want. After doing this, we would have (Beam Size)*(Beam Size + 1) candidate sequences. We divide these sequences into 2 banks, where Bank 0 is the bank with sequences which do not contain any English alphabet, and Bank 1 where the sequence satisfies the lexical constraint. We then pick the sequence from Bank 1 which has the most probability, thus satisfying our constraint. After that, we merge Banks 0 and 1 such that it is then an unconstrained problem, and we then pick the most probable sequences. Note that this means that there can be multiple sequences in the final beam sequence which satisfies the given constraint. This chain would go on till an EOT token is encountered, after which we follow the standard beam search algorithm to pick the best possible sequence. One variation could be that we only pick the most prob-

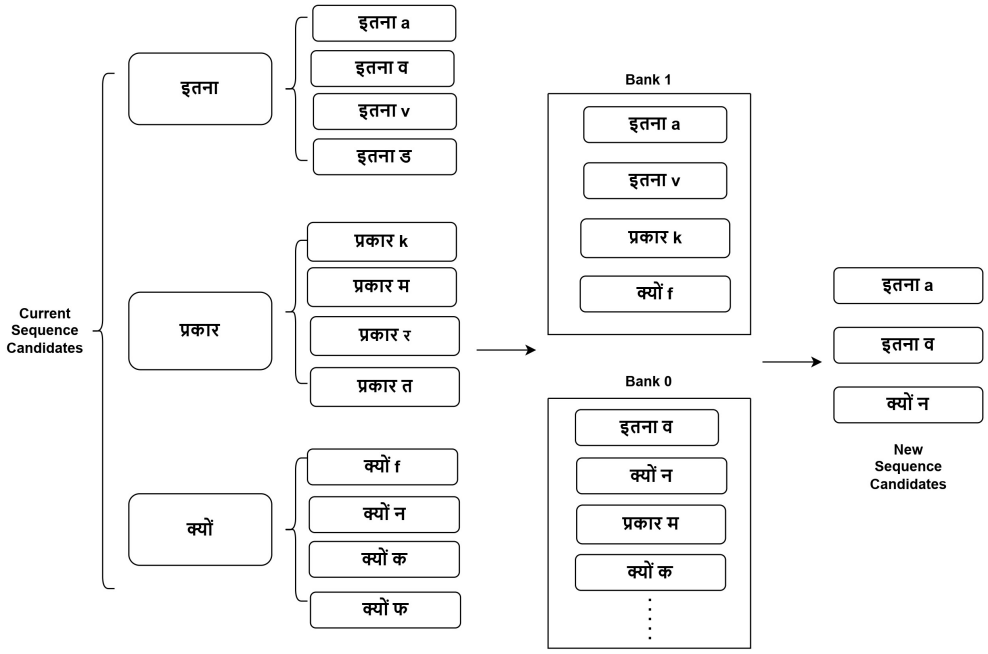

Figure 1: Diagram illustrating the Lexically Constrained Beam Search Algorithm

able sequence from the finished sequences satisfying the constraints, but we relax this condition for now. The pseudocode of the algorithm has been given in 1.

Algorithm 1 Constrained Beam Search

```
beamsize ← 3
for i in n-audio do
    for j in beamsize do
        for Top beamsize candidates do
            Sequence = previous j-sequence + new candidate
            Add Sequence to bank-1 if it has English alphabet
            Add Sequence to bank-0 if it does not have an English
alphabet
        end for
    end for
    Add best candidate from bank-1 to our top beamsize
candidates
    Merge bank-1 and bank-0 and sample beamsize − 1 candidates
from this merged pool
end for
Keep a max for final finished sequences and add EOT to
unfinished sequences if there are not enough finished
sequences
```