Given the basic architecture, the probability of any word can be written in the way shown below:

Dual Language Model

$$P[w' \mid w] = \begin{cases} P_1[w' \mid \langle \mathbf{s} \rangle] & \text{if } w' \in \mathrm{V}_1 \\ P_2[w' \mid \langle \mathbf{s} \rangle] & \text{if } w' \in \mathrm{V}_2 \\ 0 & \text{if } w' = \langle /\mathbf{s} \rangle \end{cases} \qquad \text{for } w = \langle \mathbf{s} \rangle$$

$$P[w' \mid w] = \begin{cases} P_1[w' \mid w] & \text{if } w' \in \mathrm{V}_1 \cup \{\langle/\texttt{s}\rangle\} \\ P_1[\langle\texttt{sw}\rangle \mid w] \cdot P_2[w' \mid \langle\texttt{sw}\rangle] & \text{if } w' \in \mathrm{V}_2 \end{cases} \qquad \text{for } w \in \mathrm{V}_1$$

$$P[w' \mid w] = \begin{cases} P_2[w' \mid w] & \text{if } w' \in \mathrm{V}_2 \cup \{\langle\texttt{/s}\rangle\} \\ P_2[\langle\texttt{sw}\rangle \mid w] \cdot P_1[w' \mid \langle\texttt{sw}\rangle] & \text{if } w' \in \mathrm{V}_1 \end{cases} \qquad \text{for } w \in \mathrm{V}_2$$

Dual Language Model

Given the basic architecture, the probability of any word can be written in the way shown below:

$$P[w' \mid w] = \begin{cases} P_1[w' \mid \langle \mathrm{s} \rangle] & \text{if } w' \in \mathrm{V}_1 \\ P_2[w' \mid \langle \mathrm{s} \rangle] & \text{if } w' \in \mathrm{V}_2 \\ 0 & \text{if } w' = \langle /\mathrm{s} \rangle \end{cases} \qquad \text{for } w = \langle \mathrm{s} \rangle$$

$$P[w' \mid w] = \begin{cases} P_1[w' \mid w] & \text{if } w' \in \mathrm{V}_1 \cup \{\langle /\mathrm{s} \rangle\} \\ P_1[\langle \mathtt{sw} \rangle \mid w] \cdot P_2[w' \mid \langle \mathtt{sw} \rangle] & \text{if } w' \in \mathrm{V}_2 \end{cases} \qquad \text{for } w \in \mathrm{V}_1$$

$$P[w' \mid w] = \begin{cases} P_2[w' \mid w] & \text{if } w' \in \mathrm{V}_2 \cup \{\langle /\mathrm{s} \rangle\} \\ P_2[\langle \mathtt{sw} \rangle \mid w] \cdot P_1[w' \mid \langle \mathtt{sw} \rangle] & \text{if } w' \in \mathrm{V}_1 \end{cases} \qquad \text{for } w \in \mathrm{V}_2$$

FST for DLM