

## K smallest edit distances

Let  $lev_{a,b}(i, j)$  be the set of  $k$  smallest edit distances between the first  $i$  characters of the string  $a$  and the first  $j$  characters of the string  $b$ .

Let's introduce the following sets:

$I_{a,b}(i, j) := \{d + 1 \mid d \in lev_{a,b}(i - 1, j)\}$	The set of $k$ edit distances obtained through insertions
$D_{a,b}(i, j) := \{d + 1 \mid d \in lev_{a,b}(i, j - 1)\}$	The set of $k$ edit distances obtained through deletions
$S_{a,b}(i, j) := \{d + 1_{(a_i \neq b_j)} \mid d \in lev_{a,b}(i - 1, j - 1)\}$	The set of $k$ edit distances obtained through substitutions
$E_{a,b}(i, j) := I_{a,b}(i, j) \cup D_{a,b}(i, j) \cup S_{a,b}(i, j)$	The set of all $3k$ edit distances

Where  $1_{(a_i \neq b_j)}$  is the indicator function equal to 0 when  $a_i = b_j$  and equal to 1 otherwise.

Let's consider the set  $E_{a,b,k}(i, j) \subseteq E_{a,b}(i, j)$ , such that  $|E_{a,b,k}(i, j)| = k$ , and for every  $x \in E_{a,b,k}(i, j)$  and for every  $y \in E_{a,b}(i, j) \setminus E_{a,b,k}(i, j)$  it follows that  $x < y$ .

Then the set of  $k$  smallest edit distances between the first  $i$  characters of the string  $a$  and the first  $j$  characters of the string  $b$  is defined as follows:  $lev_{a,b}(i, j) := E_{a,b,k}(i, j)$ .

Yurii Lahodiuk, 26.11.2017