

Symbol	Description
$M$	The number of structure templates retained after pruning
$L$	The maximum span of records (i.e., the maximum number of lines each record can span)
$\alpha$	The minimum coverage threshold for records
$n$	The total number of lines in the dataset
$K$	The number of structure templates retained after generation
$T_{data}$	The total size of the dataset
$S_{data}$	The amount of data sampled during all three steps
$c$	The number of special characters (i.e., characters in <i>RT-CharSet-Candidate</i> ) appearing in the dataset