The input to the $Default\_Sep$ function is the $text$ variable, which corresponds to the text of the review. This function returns 2 elements which are $words$ and $shifts$.

$words$ is a list of strings. Each element of this list is a word extracted from $text$.

$shifts$ is a list of integers. Each element of this list indicates how many characters separate the word of index i from the one of index i-1. By definition, the first element of $shifts$ is a 0.

Algorithm of $Default\_Sep$ :

$character_{split} \leftarrow [\,]$

$for\ each\ character\ C\ in\ text :$

    $if\ C\ is\ not\ alpha\ numeric :$

        $if\ C\ not\ in\ character_{split}\ and\ C\ not\ in\ [',/',\%] :$

            $add\ C\ to\ character_{split}$

$for\ each\ character\ C\ in\ character_{split} :$

    $split\ text\ by\ C$

$words \leftarrow [\,]$

$for\ each\ element\ W\ in\ the\ splitted\ text :$

    $if\ W\ is\ not\ "" :$

        $add\ W\ to\ words$


$shifts \leftarrow [\,]$

$shift, i, var_{temp} \leftarrow 0,0, ""$

$for\ each\ character\ C\ in\ text :$

    $if\ C\ not\ in\ character_{split}\ or\ C\ not\ in\ [',/',\%] :$

        $add\ C\ to\ var_{temp}$

    $else :$

        $shift \leftarrow shift + 1$

    $if\ var_{temp}\ is\ equal\ to\ the\ i - th\ word\ of\ words :$

        $add\ shift\ to\ shifts$

        $shift, var_{temp} \leftarrow 0, ""$

        $if\ i\ is\ less\ than\ the\ number\ of\ words :$

            $i \leftarrow i + 1$