Все литеральные константы должны быть приведены в вид, в котором они будут использованы в программе.

Лексер должен определять ошибки связанные с отдельными лексемами.

REGEXP для перевода строки, чтобы завести там счетчик и компилятор показывал строку в которой найдена ошибка.

%x STRING

%%

%{char buffer[1000]; %}

\” { BEGIN(STRING); buffer[0] = 0;}

<STRING>[^\”\\\n]+ {strcat(buffer, yytext)}

<STRING> [\\n](file://n) {strcat(buffer, “\n”);}

<STRING> \\\\ {strcat(buffer, “\\”);}

<STRING> \\\” {strcat(buffer, “\””);}

<STRING> [\\\x[0-9a-fA-F]{1,2}](file:///x[0-9a-fA-F]%7b1,2%7d) {char str[2]; char \*a; str[0] = strtol(yytext+2, &a, 16); str[1] = 0;

strcat(buffer,str);}

<STRING> \” { printf(“%s – found string \n“, buffer);

BEGIN(INITIAL);}

%x STRING

%%

%{char buffer[1000]; %}

<STRING>\n{printf(“ERROR”);}

<STRING> <<EOF>> {printf(“ERROR: no closed quote”);}