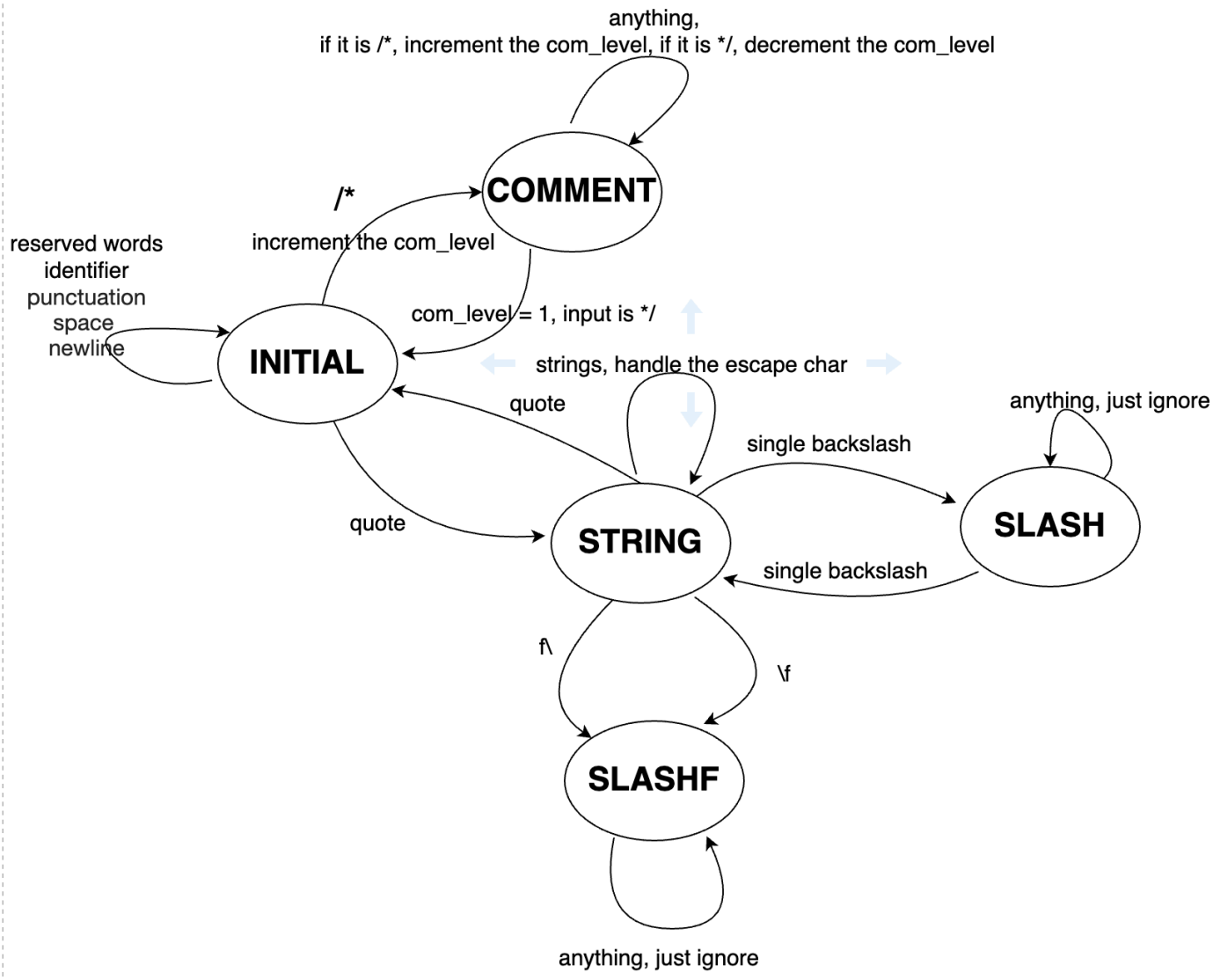
README

The figure shows the finite automata of my design:



Comments Handling

I create a state `COMMENT` and a reference `comment_level` for handling the comments.

The first time encountered the `/*`, it enter the comment state and increment the comment level. Then every time the program encounters a `/*` it just increments the comment level. If encountering a `*/`, check if the current comment level is `1`, if yes then go back to the `INITIAL` state.

String handling

Basic logic: encounter a quote, then enter the `STRING` state, encounter another quote then go back to the `INITIAL` state.

Escape Handling: There are a lot of escape characters. I handled \n, \t, \ddd, \^c , \", \\, \f … f\.

The handling logic is shown in my code and my FSM figure shown above.

Error Handling

If the regular expression can not match a char, it will raise an error, indicating the line and the pos of that character. The given files already handle that

End-of-file handling

The function of `eof` already handles the end-of-file situation. It will print out the location of the end of the file.