

---

Tokenizer For Systems Programming by Saif and Cal

This is a tokenizer written in C that can discern between words, hexadecimal, octal, floating point, and various things like brackets and other symbols. This is the first Systems Programming assignment in the class. This project was first implemented using a linked list, but was then forked so that we could remove the linked list implementation.

Steps to get it to run

- 1) Compile the source file using your favorite compiler and give the executable a name
- 2) Run the program from the command line with a command such as `./tokenizer "0x4356abdc 0777 [] "`  
or  
`tokenizer.exe "0x4356abdc 0777 [] "`

Please be sure to wrap the argument using quotes `"` otherwise the program will fail to run correctly. Certain characters such as `\` (backslash) will cause the program to fail. Backslashes are an escape character in C and in certain shell environments, please avoid using it in your inputs. Backslashes as part of a full a sentence are supported.

Emojis, and characters that do not appear on a normal keyboard are not supported and may not be tokenized.

Check the test cases for expected outputs. Most keyboard characters are supported. If our program prints "Malformed Token", you most likely entered a number with a bad format. This will trigger occasionally but not on all badly inputted numbers.