Mark Hirons mch165 Computer Architecture Russell Spring 2016

Description
---

tokenizer.c is a basic string tokenizer that parses integer tokens of the decimal, octal, and hex types as well as floating point tokens from a single string argument. This program utilizes a fsm(finite state machine) to separate tokens and determine their type.

String arguments are entered as a string between quotations in the command line after typing "./tokenizer".

tokenizer.c will print the type of any valid tokens and said token on a single line.

Features

-Only valid tokens will be printed in the form:

```
type            token
```

-Bad characters are printed in an error message of the form:

```
error:          [hexcode]
```

Bad characters are those that don't belong to a token and end a valid token. The error message will be printed on the line immediately BEFORE the valid token it ends.
See Test Case 27

-Invalid tokens will result in an error message of the form:

```
error:          mal token
```

An invalid token is one which does not fit in the category of an integer number(decimal, hex, octal forms) or a floating point number.

-The tokenizer does not alter the original string argument
-The tokenizer uses only the minimal amount of memory to store a copy of the original string argument, and to build tokens. Tokens are built character by character.

**Tokens can be be returned by themselves by calling the TKGetNextToken function. The type of the current token is returned as a separate string from the TKGetType function.