2.3.3 continued

\u sequences are processed BEFORE the code is parsed.

\u gotcha
related to
strings

"\u0022+\u0022" ⟹ ""+"" ⟶ empty string

the compiler sees
2 quotes, a plus, and two quotes
NOT a string containing a plus
surrounded by double quotes

Comments aren't safe either

\u gotcha
related to
comments

    //   \u000A here is text

newline, so compiler sees "here is text" on the next line,
uncommented, so :

    //   \u000A methodA();

this method will execute even though it looks like
it's in a comment

Also note:

    // ~~c:\windows~~ C:\users

this looks like a unicode escape, but not followed
by hex digits, so syntax error.

multiple u's ok
\uuCAFE
= \uCAFE

Can have any number of u in \u, e.g. \uuu000A means same thing.
Done to allow conversion tools to be able to reverse conversions to get exact
source back.

exercise: write a program that demonstrates these gotchas.