DragonBook

Question 1.1.4: A compiler that translates a high-level language into another high-level language is called a source-to-source translator. What advantages are there to using C as a target language for a compiler?

- The reason C would be a good target language for a compiler is because C is an ultimately easy to understand and in turn easy to generate a compiler for. Another reason is because C compilers can aggressively optimize meaning that you don't have to worry about register allocation, common sub-expression elimination, or loop unrolling.

Question 1.6.1: For the block-structured C code of Fig. 1.13(a), indicate the values assigned to w, x, y, and z

- W = 13
- X = 11
- Y = 13
- Z = 11

Crafting a Compiler

Question 1.11: Investigate the techniques MOSS uses to find similarity. How does MOSS differ from other approaches for detecting possible plagiarism?

- The way that MOSS works is that it relies on a document fingerprinting algorithm called winnowing. The way winnowing works is that it takes two documents and computes each's fingerprint and queries minute with a precomputed index of fingerprints for all documents of a reference collection. Minute matching can indicate shared text segments and suggest potential plagiarism. Other methods differ because they use different approaches such as citation analysis which depends solely on citation and reference information rather than on textual similarity or even Stylometry which quantifies an authors unique writing style.

Question 3.1: What token sequence is produced? For which tokens must extra information be returned in addition to the token code?

<const> <type, float> <id, payment> <assign> <payment, 384.00> <semi>

<type, float> <id, bal> <semi>

<type, int> <id, month> <assign> <month, 0> <semi>

<id, bal> <assign> <bal , 15000> <semi>

<**while**> <leftparen> <id, bal> <lessthan> <digit, 0> <right paren> <leftbracket>

<**printf**> <leftparen> <quot> <string> < comma > <id, month> <id, bal> <rightparen> <semi>

<id, bal> <assign> <id, bal> <minus> <id, payment> <plus><float, 0.015 <times><id, bal> <semi>

<id, month> <assign> <id, month> <plus> <digit, 1> <semi>

<rightbracket>

<rightbracket>

More information must be returned for tokens that hold more information than regular token such as id's, types, and ints.