Compiler Nots

This compilation highlights key points from each chapter that I deem essential for memorization. Following each chapter, you'll find solutions to the challenges presented, providing a comprehensive understanding of the material.

`Status` `Processing`

References

- Crafting Interpreters book
- Crafting Interpreters Repo

Table of Contents

Chapter 1 Introduction

NOTES

1. little languages = domain-specific languages : These are pidgins tailor-built to a specific task.
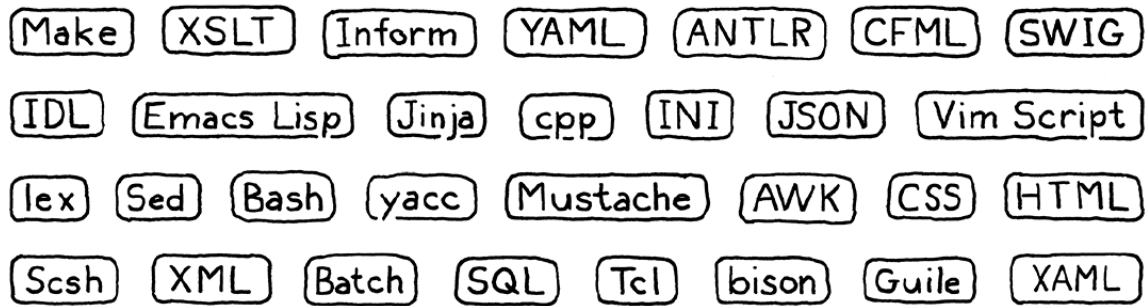
Figure 1: domain-specific languages

2. A compiler reads files in one language. translates them, and outputs files in another language. You can implement a compiler in any language, including the same language it compiles, a process called *self-hosting*.

3. You can't compile your compiler using itself yet, but if you have another compiler for your language written in some other language, you use that one to compile your compiler once. Now you can use the compiled version of your own compiler to compile future versions of itself and you can discard the original one compiled from the other compiler. This is called *bootstrapping* from the image of pulling yourself up by your own bootstraps.

4. Bytecode is an intermediate representation (IR) of a program's source code that is generated by a compiler or an interpreter before it is executed. It is a low-level, platform-independent set of instructions that can be executed by a virtual machine (VM) or interpreter. Bytecode is often used to bridge the gap between high-level programming languages and machine code.

CHALLENGES

1. There are at least six domain-specific languages used in the little system I cobbled together to write and publish this book. What are they?

   > HTML , CSS , SQL , JSON , XML , XAML , Bash

2. Get a "Hello, world!" program written and running in Java. Set up whatever Makefiles or IDE projects you need to get it working. If you have a debugger, get comfortable with it and step through your program as it runs.

   > Chapter1.java

   ```
   1  public class Chapter1 {
   2    public static void main(String[] args) {
   3        System.out.println("Hello, World!");
   ```

```
4   }
5 }
```

Compile the program with the following command:

```
1  javac Chapter1.java
```

Run the compiled program with:

```
1   java Chapter1
```

3. Do the same thing for C. To get some practice with pointers, define a doubly-linked list of heap-allocated strings. Write functions to insert, find, and delete items from it. Test them.

doubly-linked-list.c

```c
1  #include <stdio.h>
2  #include <stdlib.h>
3  #include <string.h>
4
5  typedef struct Node {
6      char* data;
7      struct Node* prev;
8      struct Node* next;
9  } Node;
10
11 typedef struct {
12     Node* head;
13     Node* tail;
14     int size;
15 } LinkedList;
16
17 void initList (LinkedList *list){
18     list->head = NULL;
19     list->tail = NULL;
20     list->size = 0;
21 }
22
23 void insert(LinkedList *list , const char* data) {
24     Node *newNode = (Node*) malloc(sizeof(Node));
25     newNode->data = strdup(data);
26     newNode->prev = list->tail; // list->tail to add the new node
            at end
27     newNode->next = NULL; // now the new node at the end of the
            list
28     if(list->tail != NULL) {
29         list->tail->next = newNode;
30     }
```

```c
31      list->tail = newNode;
32      if(list->head == NULL) {
33          list->head = newNode;
34      }
35      list->size++;
36  }
37
38  Node *find (LinkedList *list , const char* data) {
39      Node *cuur = list->head;
40      while (cuur != NULL) {
41          if (strcmp(cuur->data, data) == 0) {
42              return cuur;
43          }
44          cuur = cuur->next;
45      }
46      return NULL;
47  }
48
49  void deleteNode(LinkedList *list, Node *node) {
50      if (node == NULL) {
51          return;
52      }
53      if(node->prev != NULL) {
54          node->prev->next = node->next;
55      } else {
56          list->head = node->next;
57      }
58      if(node->next != NULL) {
59          node->next->prev = node->prev;
60      } else {
61          list->tail = node->prev;
62      }
63
64      free(node->data);
65      free(node);
66      list->size--;
67  }
68
69  void freeList (LinkedList *list) {
70      Node *curr = list->head;
71      while (curr != NULL) {
72          Node *next = curr->next;
73          free(curr->data);
74          free(curr);
75          curr = next;
76      }
77      list->head = NULL;
78      list->tail = NULL;
79      list->size = 0;
80  }
81
```

```
82  int getSize (LinkedList *list) {
83      return list->size;
84  }
85
86  void printList (LinkedList *list) {
87      Node *cuur = list->head;
88      while (cuur != NULL) {
89          printf("%s ",cuur->data);
90          cuur = cuur->next;
91      }
92      printf("\n");
93  }
94
95  int main () {
96      LinkedList list;
97      initList(&list);
98      insert(&list, "Hello");
99      insert(&list, "Hussein");
100     insert(&list, "Hussein2");
101
102     printf("Original List: ");
103     printList(&list);
104
105     int size = getSize(&list);
106     printf("Size: %d\n" , size);
107
108     Node *foundNode = find(&list, "Hussein2");
109     if (foundNode != NULL) {
110         printf("Found: %s\n", foundNode->data);
111         deleteNode(&list, foundNode);
112         printf("After Deletion: ");
113         printList(&list);
114     } else {
115         printf("Not Found.\n");
116     }
117
118     freeList(&list);
119
120     return 0;
121 }
```

Compile the program with the following command:

```
1  gcc doubly-linked-list.c
```

Run the compiled program with:

```
1  .\doubly-linked-list
```

Chapter 2 A Map of the Territory

NOTES

1. This book is about a language's implementation, which is distinct from the language itself in some sort of Platonic ideal form. Things like *stack*, *bytecode*, and *recursive descent*, are nuts and bolts one particular implementation might use.
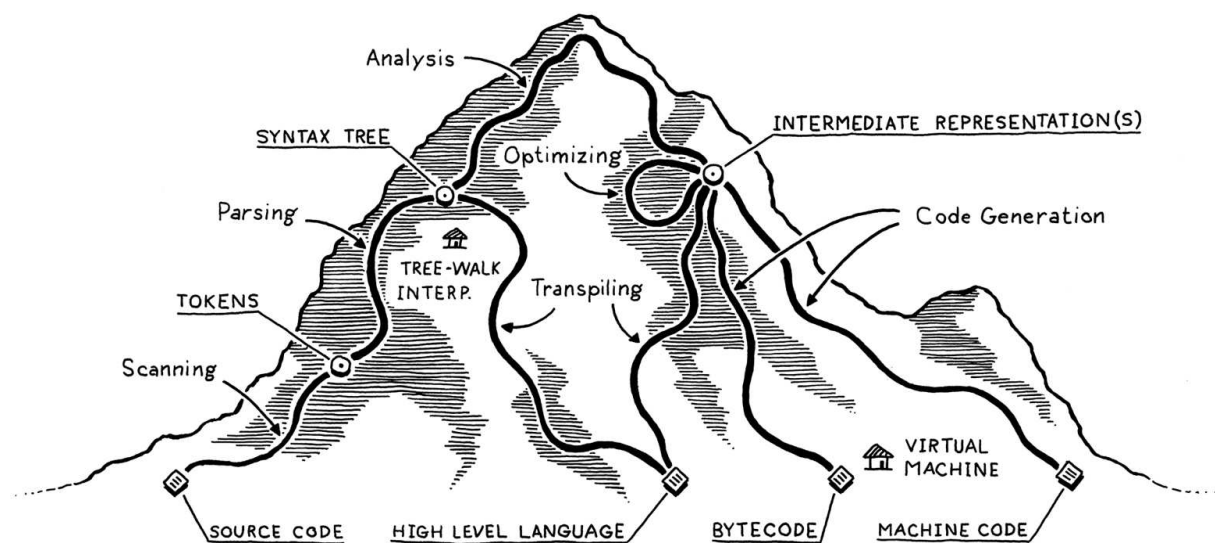

Figure 2: overall view

2. A scanner (or lexer) takes in the linear stream of characters and chunks them together into a series of something more akin to *words*.

3. A parser takes the flat sequence of tokens and builds a tree structure that mirrors the nested nature of the grammar. These trees have a couple of different names—*parse tree* or *abstract syntax tree* —depending on how close to the bare syntactic structure of the source language they are. In practice, language hackers usually call them *syntax trees*, *ASTs*, or often just *trees*. The parser's job also includes letting us know when we do by reporting *syntax errors*.

4. The language we'll build in this book (lox) is dynamically typed, so it will do its type checking later, at runtime.

5. The front end of the pipeline is specific to the source language the program is written in. The back end is concerned with the final architecture where the program will run.

6. Intermediate representation lets you support multiple source languages and target platforms with less effort. Say you want to implement *Pascal*, *C* and *Fortran* compilers and you want to target *x86*, *ARM*, and, I dunno, *SPARC*. Normally, that means you're signing up to write **nine full**

compilers: Pascal→x86, C→ARM, and every other combination. A shared intermediate representation reduces that dramatically. You write one front end for each source language that produces the IR. Then one back end for each target architecture. Now you can mix and match those to get every combination.

- Source Languages: C, Fortran, Pascal

- Target Architectures: x86, ARM, SPARC

Without Intermediate Representation (IR)

1. C

  - x86
  - ARM
  - SPARC

2. Fortran

  - x86
  - ARM
  - SPARC

3. Pascal

  - x86
  - ARM
  - SPARC

> 9 full compilers

With Intermediate Representation (IR)

Front-end

1. C

  - IR

2. Fortran

  - IR

3. Pascal

  - IR

Back-end

4. IR

   – x86
   – ARM
   – SPARC

   > 3 full compilers

7. If we generate real *machine code*, we get an executable that the OS can load directly onto the chip. *Native code* is lightning fast, but generating it is a lot of work. Today's architectures have piles of instructions, complex pipelines, and enough historical baggage to fill a 747's luggage bay.

8. **Execution Environments:**

   • **JVM:**
     – Executes bytecode.

   • **.Net:**
     – Uses CLR (Common Language Runtime).
     – Executes IL (Microsoft Intermediate Language).

9. **Bytecode Compilation Options:**

   If your compiler produces bytecode, you have two primary options:

   • **Option 1: Native Code Compilation**

     – Write a mini-compiler for each target architecture.
     – Convert the bytecode to native code for the specific machine.

   • **Option 2: Virtual Machine (VM) Execution**

     – Write a virtual machine (VM) program.
     – The VM emulates a hypothetical chip supporting your virtual architecture at runtime.
     – Running bytecode in a VM is slower than translating it to native code ahead of time.
     – Offers simplicity and portability.
     – Implement the VM in a language like C, allowing the language to run on any platform with a C compiler.
     – This approach is employed by the second interpreter developed in this book.

10. If the language is run inside an interpreter or VM, then the runtime lives there. This is how most implementations of languages like Java, Python, and JavaScript work.

11. *Syntax-directed translation* is a structured technique for building these all-at-once compilers. You associate an action with each piece of the grammar, usually one that generates output code. Then, whenever the parser matches that chunk of syntax, it executes the action, building up the target code one rule at a time.

12. The fastest way to execute code is by compiling it to machine code, but you might not know what architecture your end user's machine supports. What to do? > You can do the same thing that the HotSpot *JVM*, *Microsoft's CLR* and most *JavaScript interpreters* do. On the end user's machine, when the program is loaded—either from source in the case of JS, or platform-independent bytecode for the JVM and CLR—you compile it to *native* for the architecture their computer supports. Naturally enough, this is called *just-in-time compilation*. Most hackers just say "JIT".

13. Compilers and Interpreters

    - *Compiling* is an implementation technique that involves translating a source language to some other—usually lower-level—form. When you generate bytecode or machine code, you are compiling. When you transpile to another high-level language you are compiling too.
    - When we say a language implementation *is a compiler*, we mean it translates source code to some other form but doesn't execute it.
    - when we say an implementation "is an interpreter", we mean it takes in source code and executes it immediately. it runs programs "from source".

14. CPython is an interpreter, and it has a compiler

CHALLENGES

1. Pick an open source implementation of a language you like. Download the source code and poke around in it. Try to find the code that implements the scanner and parser. Are they hand-written, or generated using tools like Lex and Yacc? (.l or .y files usually imply the latter.)

2. Just-in-time compilation tends to be the fastest way to implement a dynamically typed language, but not all of them use it. What reasons are there to not JIT?

   > Complexity, Portability, and Compilation Overhead.

3. Most Lisp implementations that compile to C also contain an interpreter that lets them execute Lisp code on the fly as well. Why?

Chapter 3 The Lox Language

NOTES

1. Lox is dynamically typed.

2. Automatic memory management

3. There are two main techniques for managing memory: reference counting and tracing garbage collection (usually just called "garbage collection" or "GC").

4. Data Types:

Booleans

```
1  true; // Not false.
2  false; // Not *not* false.
```

Numbers: Lox only has one kind of number: double-precision floating point.

```
1  1234; // An integer.
2  12.34; // A decimal number.
3
4  .2; // not allowed in lox
5  2.; // not allowed in lox
```

Strings

```
1  "I am a string";
2  ""; // The empty string.
3  "123"; // This is a string, not a number.
```

Nil

```
1  return nil; // similar to returning null in other languages
```

5. Expressions:

Arithmetic

```
1  add + me;
2  subtract - me;
3  multiply * me;
4  divide / me;
5
6  -negateMe;
```

Comparison and equality : 0 in lox is true not false

```
1   less < than;
2   lessThan <= orEqual;
3   greater > than;
4   greaterThan >= orEqual;
5
6   1 == 2; // false.
7   "cat" != "dog"; // true.
8
9   314 == "pi"; // false.
10
11  123 == "123"; // false.
```

- look at this function in the interpreter

```
1    // the 0 in lox is true not false, if U want it be false edit the
         function below
2    // < check-operands
3    // > is-truthy
4    private boolean isTruthy(Object object) {
5      if (object == null)
6        return false;
7      if (object instanceof Boolean)
8        return (boolean) object;
9      return true;
10   }
```

Logical operators : The reason and and or are like control flow structures is because they short circuit. Not only does and return the left operand if it is false, it doesn't even evaluate the right one in that case.

```
1  !true; // false.
2  !false; // true.
3
4  true and false; // false.
5  true and true; // true.
6
7  false or false; // false.
8  true or false; // true.
```

```
1  var average = (min + max) / 2;
```

6. Statements

```
1  print "Hello, world!";
2
3  "some expression";
4
5  {
6  print "One statement.";
7  print "Two statements.";
8  }
```

7. Variables

```
1  var imAVariable = "here is my value";
2  var iAmNil;
3
4  var breakfast = "bagels";
5  print breakfast; // "bagels".
6  breakfast = "beignets";
7  print breakfast; // "beignets".
```

8. Control Flow

```
1  if (condition) {
2  print "yes";
3  } else {
4  print "no";
5  }
6
7  var a = 1;
8  while (a < 10) {
9  print a;
10  a = a + 1;
11  }
12
13  for (var a = 1; a < 10; a = a + 1) {
14  print a;
15  }
```

9. Functions

- An *argument* is an actual value you pass to a function when you call it.

- A *parameter* is a variable that holds the value of the argument inside the body of the function.

```
1  makeBreakfast(bacon, eggs, toast);
2
3  makeBreakfast();
4
5  // a and b called parameters
6  fun printSum(a, b) {
7      print a + b;
8  }
9
10 // 1 and 2 called arguments
11 printSum(1,2);
```

Closures

```
1  fun addPair(a, b) {
2  return a + b;
3  }
4
5  fun identity(a) {
6  return a;
7  }
8  print identity(addPair)(1, 2); // Prints "3".
9
10 fun outerFunction() {
11     fun localFunction() {
12         print "I'm local!";
13     }
14     localFunction();
15 }
16
17 fun returnFunction() {
18     var outside = "outside";
19     fun inner() {
20     print outside;
21     }
22     return inner;
23 }
24 var fn = returnFunction();
25 fn();
```

10. Classes

```
1  class Breakfast {
2      // var x = 5; // this is not allowed in lox
3      cook() {
4          print "Eggs a-fryin'!";
5      }
6      serve(who) {
7          print "Enjoy your breakfast, " + who + ".";
```

```
 8        }
 9  }
10
11  // Store it in variables.
12  var someVariable = Breakfast;
13  // Pass it to functions.
14  someFunction(Breakfast);
15
16  var breakfast = Breakfast();
17  print breakfast; // "Breakfast instance".
```

Inheritance: using a less-than (<) operator

```
 1  class Brunch < Breakfast {
 2      drink() {
 3          print "How about a Bloody Mary?";
 4      }
 5  }
 6
 7  var benedict = Brunch("ham", "English muffin");
 8  benedict.serve("Noble Reader");
 9
10  class Brunch < Breakfast {
11      init(meat, bread, drink) {
12          super.init(meat, bread);
13          this.drink = drink;
14      }
15  }
```

The Standard Library : built-in function clock() that returns the number of seconds since the program started.

CHALLENGES

1. Write some sample Lox programs and run them (you can use the implementations of Lox in my repository). Try to come up with edge case behavior I didn't specify here. Does it do what you expect? Why or why not?

```
 1  /*
 2  fun makeCounter() {
 3    var i = 0;
 4    fun count() {
 5      i = i + 1;
 6      print i;
 7    }
 8
 9    return count;
```

```
10  }
11
12  //fun scope(a) {
13  //   print a; // parameter
14  //   var a = "local";
15  //   print a; // local
16  //}
17
18  fun thrice(fn) {
19    for (var i = 1; i <= 3; i = i + 1) {
20      fn(i);
21    }
22  }
23
24  thrice(fun (a) {
25    print a;
26  });
27  // "1".
28  // "2".
29  // "3".
30
31  var counter = makeCounter();
32  //counter(); // "1".
33  //counter(); // "2".
34
35  //scope("parameter");
36
37  fun scope(a) {
38    print a;
39    var a = "local";
40    print a;
41  }
42
43  print 10;
44  scope(5);
45
46  /*
47
48  /*
49  var a = "global";
50  {
51    fun showA() {
52      print a;
53    }
54
55    showA();
56    var a = "block";
57    showA();
58    {
59      var a = "block2";
60      showA();
```

```
 61      }
 62      showA();
 63  }
 64  */
 65
 66  /*
 67  var a = 5;
 68
 69  {
 70    print a;
 71    var a = a;
 72    print a;
 73    a = 6;
 74    print a;
 75  }
 76  */
 77
 78
 79
 80  /*
 81  var a = 5;
 82  var a = 6;
 83  print a;
 84  */
 85
 86
 87
 88  /*
 89  fun bad() {
 90  var a = "first";
 91  var a = "second";
 92  print a;
 93  }
 94
 95  bad();
 96  */
 97
 98  //break;
 99
100  /*
101  while (true) {
102    if (5 > 0 ) {
103      print 6;
104      break;
105    }
106    var a = 6;
107    print a;
108  }
109  var a = 5;
110  print a;
111  */
```

```
112
113  /*
114  while (true) {
115    var a = 6;
116    print a;
117
118    if (5 > 0) {
119      break;
120    }
121  }
122
123  var a = 5;
124  print a;
125  */
126
127  /*
128  if (a > 1) {
129    print a;
130    break;
131  }
132  */
133
134  /*
135  class DevonshireCream {
136    serveOn() {
137      return "Scones";
138    }
139  }
140  print DevonshireCream; // Prints "DevonshireCream".
141
142  class Bagel {}
143  var bagel = Bagel();
144  print bagel; // Prints "Bagel instance".
145
146  class Bacon {
147    eat() {
148      print "Crunch crunch crunch!";
149    }
150  }
151
152  Bacon().eat(); // Prints "Crunch crunch crunch!".
153  */
154
155  /*
156  class Doughnut {
157    cook() {
158      print "Fry until golden brown.";
159    }
160  }
161  class BostonCream < Doughnut {
162    cook() {
```

```
163       var method = super.cook;
164       method();
165       super.cook();
166       print "Pipe full of custard and coat with chocolate.";
167     }
168   }
169   BostonCream().cook();
170   */
171
172   /*
173   class A {
174     method() {
175     print "Method A";
176     }
177   }
178   class B < A {
179     method() {
180       print "Method B";
181     }
182     test() {
183       super.method();
184     }
185   }
186   class C < B {}
187
188   C().test();
189   */
190
191   /*
192   fun fib(n) {
193     if (n < 2) return n;
194     return fib(n - 1) + fib(n - 2);
195   }
196
197   var before = clock();
198   print fib(40);
199   var after = clock();
200   print after - before;
201   */
202
203   while (true) {
204     if (5 > 0 ) {
205       print 6;
206       break;
207     }
208   }
```

2. This informal introduction leaves a lot unspecified. List several open questions you have about the language's syntax and semantics. What do you think the answers should be?

3. Lox is a pretty tiny language. What features do you think it is missing that would make it annoy-

ing to use for real programs? (Aside from the standard library, of course.)

Chapter 4 Scanning

NOTES

1. The scanner takes in raw source code as a series of characters and groups it into a series of chunks we call tokens.

2. Lox is a scripting -high level- language, which means it executes directly from source.

3. Each of these blobs of characters is called a lexeme:

Figure 3: blobs of characters (lexeme)

4. Note that the ! and = are not two independent operators. You can't write ! = in Lox and have it behave like an inequality operator.

5. We've got another helper:

```
1  private char peek() {
2    if (isAtEnd()) return '\0';
3    return source.charAt(current);
4  }
```

It's sort of like advance(), but doesn't consume the character. This is called **_lookahead_**. Since it only looks at the current unconsumed character, we have one character of lookahead. **_The smaller this number is, generally, the faster the scanner runs._** The rules of the lexical grammar dictate how much lookahead we need. Fortunately, most languages in wide use only peek one or two characters ahead.

Mid Qs : What is the lookahead of the scanner ? is it better to have larger or smaller lookaheads ? What is the lookahead of Lox ?

- Lookahead in Scanners:

Meaning: The lookahead of a scanner refers to the number of characters it can examine ahead of its current position without actually consuming them. It aids in making informed decisions about token formation.

- Optimal Size:

Smaller lookaheads (1-2 characters) are generally preferred for efficiency. Larger lookaheads might be necessary for certain language constructs but can impact speed.

- Lox's Lookahead

Lox's scanner uses a lookahead of 1 character. This is sufficient for its lexical grammar, as it doesn't have complex constructs that require extensive lookahead.

6. Since we only look for a digit to start a number, that means -123 is not a number literal. Instead, -123, is an expression that applies - to the number literal 123.

```
1  print -123.abs();
```

This prints -123 because negation has lower precedence than method calls. We could fix that by making - part of the number literal. But then consider:

```
1  var n = 123;
2  print - n.abs();
```

This still produces -123, so now the language seems inconsistent. No matter what you do, some case ends up weird.

7. We don't allow a leading or trailing decimal point, so these are both invalid:

```
1  .1234
2  1234.
```

8. We don't allow this :

```
1  123.sqrt();
```

9. Consider this in lox:

```
1  var a = 5;
2
3  print -a; // -5
4  print --a; // 5
5  print ---a; // -5
```

10. `java case 'o': if (peek()== 'r'){ addToken(OR); } break;`

Consider what would happen if a user named a variable *orchid*. The scanner would see the first two letters, or, and immediately emit an or keyword token. This gets us to an important principle called **_maximal munch_**. *When two lexical grammar rules can both match a chunk of code that the scanner is looking at, whichever one matches the most characters wins.*

11. Maximal munch means we can't easily detect a reserved word until we've reached the end of what might instead be an identifier.

CHALLENGES

1. The lexical grammars of Python and Haskell are not regular. What does that mean, and why aren't they?

   > A lexical grammar defines the basic building blocks of a programming language, such as tokens (keywords, identifiers, literals, etc.), and specifies how these tokens are combined to form valid programs. The reason why the lexical grammars of Python and Haskell are not regular can be attributed to the expressive power and flexibility that these languages provide.

2. Aside from separating tokens—distinguishing print foo from printfoo — spaces aren't used for much in most languages. However, in a couple of dark corners, a space does affect how code is parsed in CoffeeScript, Ruby, and the C preprocessor. Where and what effect does it have in each of those languages?

3. Our scanner here, like most, discards comments and whitespace since those aren't needed by the parser. Why might you want to write a scanner that does not discard those? What would it be useful for?

4. Add support to Lox's scanner for C-style /*….. */ block comments. Make sure to handle newlines in them. Consider allowing them to nest. Is adding support for nesting more work than you expected? Why?

- Scanner.java

```
1  case '/':
2                  if (match('/')) {
3                      // A comment goes until the end of the line.
4                      while (peek() != '\n' && !isAtEnd())
5                          advance();
6                  } else if (match('*')) {
7                      MultilineComment();
8                  } else {
9                      addToken(SLASH);
10                 }
11                 break;
```

```
1      private void MultilineComment() {
2
3          int nestLevel = 1;
```

```
 4
 5          while (true) {
 6              if (isAtEnd()) {
 7                  Lox.error(line, "Unterminated multiline comment.");
 8                  return;
 9              }
10
11              if (peek() == '/' && peekNext() == '*') {
12                  // Consume the '*' and '/' characters.
13                  advance();
14                  advance();
15                  nestLevel++;
16                  return;
17              } else if (peek() == '*' && peekNext() == '/') {
18                  advance();
19                  advance();
20                  nestLevel--;
21
22                  if (nestLevel == 0) {
23                      return;
24
25                  }
26              } else if (peek() == '\n') {
27                  line++;
28              }
29
30              advance();
31          }
32      }
```

Chapter 5 Representing Code

NOTS

1. If you start with the rules, you can use them to generate strings that are in the grammar. Strings created this way are called derivations because each is "derived" from the rules of the grammar.

2. Rules are called productions because they produce strings in the grammar.

3. A terminal is a letter from the grammar's alphabet.

4. A nonterminal is a named reference to another rule in the grammar.

```
expression      → literal
                | unary
                | binary
                | grouping ;

literal         → NUMBER | STRING | "true" | "false" | "nil" ;
grouping        → "(" expression ")" ;
unary           → ( "-" | "!" ) expression ;
binary          → expression operator expression ;
operator        → "==" | "!=" | "<" | "<=" | ">" | ">="
                | "+"  | "-"  | "*" | "/" ;
```

Figure 4: grammar

5. To perform an operation on a pastry, we call its accept() method and pass in the visitor for the operation we want to execute. The pastry—the specific subclass's overriding implementation of accept()—turns around and calls the appropriate visit method on the visitor and passes itself to it. That's the heart of the trick right there._ **It lets us use polymorphic dispatch on the pastry classes to select the appropriate method on the visitor class._**

CHALLENGES

1. Earlier, I said that the |, *, and + forms we added to our grammar metasyntax were just syntactic sugar. Given this grammar:

   expr → expr ("(" (expr ("," expr)*)? ")" | "." IDENTIFIER)+ | IDENTIFIER | NUMBER

Produce a grammar that matches the same language but does not use any of that notational sugar. Bonus: What kind of expression does this bit of grammar encode?

2. The Visitor pattern lets you emulate the functional style in an object-oriented language. Devise a complementary pattern for a functional language. It should let you bundle all of the operations on one type together and let you define new types easily. (SML or Haskell would be ideal for this exercise, but Scheme or another Lisp works as well.)

3. In Reverse Polish Notation (RPN), the operands to an arithmetic operator are both placed before the operator, so 1 + 2 becomes 1 2 +. Evaluation proceeds from left to right. Numbers are pushed onto an implicit stack. An arithmetic operator pops the top two numbers, performs the operation, and pushes the result. Thus, this:

```
1  (1 + 2) * (4 - 3)
```

in RPN becomes:

```
1  1 2 + 4 3 - *
```

Define a visitor class for our syntax tree classes that takes an expression, converts it to RPN, and returns the resulting string.

> change the visitBinaryExpr function in AstPrinter.java from this one :

```java
1    @Override
2    public String visitBinaryExpr(Expr.Binary expr) {
3      return parenthesize(expr.operator.lexeme,
4          expr.left, expr.right);
5    }
6
7      private String parenthesize(String name, Expr... exprs) {
8      StringBuilder builder = new StringBuilder();
9
10     builder.append("(").append(name);
11     for (Expr expr : exprs) {
12       builder.append(" ");
13       builder.append(expr.accept(this));
14     }
15     builder.append(")");
16
17     return builder.toString();
18   }
```

> to this one :

```java
1      @Override
2
3    public String visitBinaryExpr(Expr.Binary expr) {
4        String left = expr.left.accept(this);
5        String right = expr.right.accept(this);
6        return left + " " + right + " " + expr.operator.lexeme;
7    }
```

Chapter 6 Parsing Expressions

NOTS

1. Precedence determines which operator is evaluated first in an expression containing a mixture of different operators. Precedence rules tell us that we evaluate the / before the -. Operators

with higher precedence are evaluated before operators with lower precedence. Equivalently, higher precedence operators are said to "bind tighter".

2. Associativity determines which operator is evaluated first in a series of the same operator. When an operator is left-associative (think "left-to-right"), operators on the left evaluate before those on the right. Since - is left associative, this expression:

```
1   5 - 3 - 1
```

is equivalent to:

```
1   (5 - 3) - 1
```

Assignment, on the other hand, is *right-associative*. This:

```
1   a = b = c
```

is equivalent to:

```
1   a = (b = c)
```

Name	Operators	Associates
Equality	== !=	Left
Comparison	> >= < <=	Left
Term	- +	Left
Factor	/ *	Left
Unary	! -	Right

Figure 5: Associates

We fix that by stratifying the grammar. We define a separate rule for each precedence level.

```
1   expression → ...
2   equality → ...
3   comparison → ...
4   term → ...
5   factor → ...
6   unary → ...
7   primary → ...
```

left-recursive problem:

```
1   factor → factor ( "/" | "*" ) unary | unary ;
```

This rule is correct, but not optimal for how we intend to parse it. Instead of a left recursive rule, we'll use a different one.

```
1  factor → unary ( ( "/" | "*" ) unary )* ;
```

3. New Grammar:

```
1  expression → equality ;
2  equality → comparison ( ( "!=" | "==" ) comparison )* ;
3  comparison → term ( ( ">" | ">=" | "<" | "<=" ) term )* ;
4  term → factor ( ( "-" | "+" ) factor )* ;
5  factor → unary ( ( "/" | "*" ) unary )* ;
6  unary → ( "!" | "-" ) unary | primary ;
7  primary → NUMBER | STRING | "true" | "false" | "nil" | "(" expression "
      )" ;
```

This grammar is more complex than the one we had before, but in return we have eliminated the previous one's ambiguity. It's just what we need to make a parser.

4. What parsing algorithm is used for Lox ?

> Recursive Descent, It's called "recursive descent" because it walks down the grammar.

5. In a top-down parser, you reach the lowest-precedence expressions first because they may in
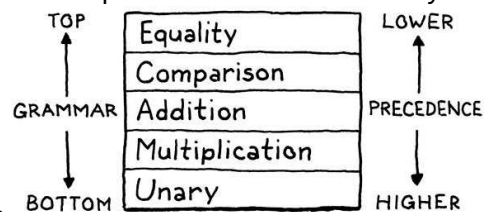
turn contain subexpressions of higher precedence.

6. Recursive descent is the simplest way to build a parser, and doesn't require using complex parser generator tools like Yacc, Bison or ANTLR.

7. Recursive descent parsers are fast, robust, and can support sophisticated errorhandling. In fact, GCC, V8 (the JavaScript VM in Chrome), Roslyn (the C# compiler written in C#) and many other heavyweight production language implementations use recursive descent.

8. Recursive descent is considered a top-down parser because it starts from the top or outermost grammar rule (here expression) and works its way down into the nested subexpressions before finally reaching the leaves of the syntax tree. This is in contrast with bottom-up parsers like LR that start with primary expressions and compose them into larger and larger chunks of syntax.

9. A recursive descent parser is a literal translation of the grammar's rules straight into imperative code. Each rule becomes a function.

10. The "recursive" part of recursive descent is because when a grammar rule refers to itself—directly or indirectly—that translates to a recursive function call.

11. A parser really has two jobs:

    1. Given a valid sequence of tokens, produce a corresponding syntax tree.
    2. Given an invalid sequence of tokens, detect any errors and tell the user about their mistakes.

12. There are a couple of hard requirements for when the parser runs into a syntax error:

    1. It must detect and report the error. If it doesn't detect the error and passes the resulting malformed syntax tree on to the interpreter, all manner of horrors may be summoned.
    2. It must not crash or hang. Syntax errors are a fact of life and language tools have to be robust in the face of them. Segfaulting or getting stuck in an infinite loop isn't allowed. While the source may not be valid code, it's still a valid input to the parser because users use the parser to learn what syntax is allowed.

13. A decent parser should:

    1. Be fast.
    2. Report as many distinct errors as there are.
    3. Minimize cascaded errors.

14. Of all the recovery techniques devised in yesteryear, the one that best stood the test of time is called—somewhat alarmingly—panic mode. As soon as the parser detects an error, it enters panic mode. It knows at least one token doesn't make sense given its current state in the middle of some stack of grammar productions. Before it can get back to parsing, it needs to get its state and the sequence of forthcoming tokens aligned such that the next token does match the rule being parsed. This process is called **synchronization**. To do that, we select some rule in the grammar that will mark the synchronization point. The parser fixes its parsing state by jumping out of any nested productions until it gets back to that rule. Then it synchronizes the token stream by discarding tokens until it reaches one that can appear at that point in the rule.

CHALLENGES

1. In C, a block is a statement form that allows you to pack a series of statements where a single one is expected. The comma operator is an analogous syntax for expressions. A comma-separated series of expressions can be given where a single expression is expected (except inside a function call's argument list). At runtime, the comma operator evaluates the le! operand and discards the result. Then it evaluates and returns the right operand. Add support for comma ex-

pressions. Give them the same precedence and associativity as in C. Write the grammar, and then implement the necessary parsing code.

> Grammar for Comma Expressions:

```
1  expression -> equality ;
2  equality -> comma ;
3  comma -> comparison ( "," comparison ) ;
```

```
1    private Expr comma() {
2      Expr expr = comparison();
3
4      while (match(COMMA)) {
5        Token operator = previous();
6        Expr right = comparison();
7        expr = new Expr.Binary(expr, operator, right);
8      }
9
10     return expr;
11   }
```

2. Likewise, add support for the C-style conditional or "ternary" operator ?:. What precedence level is allowed between the ? and :? Is the whole operator left associative or right-associative?

> Grammar for Ternary Operator:

```
1  expression -> conditional ;
2  conditional -> equality ;
```

> add this class to GenerateAst : "Conditional : Expr condition, Expr thenBranch, Expr elseBranch"

> add this function to AstPrinter

```
1    @Override
2    public String visitConditionalExpr(Expr.Conditional expr) {
3      return parenthesize2("?", expr.condition, ":", expr.thenBranch,
         expr.elseBranch);
4    }
```

> update the scanner and the need operators

> in Parser.java

```
1    private Expr conditional() {
2      Expr expr = equality();
3
4      if (match(QUESTION)) {
5        Expr thenBranch = expression();
6        consume(COLON, "Expect ':' after 'then' branch in conditional
            expression.");
7        Expr elseBranch = conditional();
8        expr = new Expr.Conditional(expr, thenBranch, elseBranch);
9      }
10
11     return expr;
12   }
```

3. Add error productions to handle each binary operator appearing without a left hand operand. In other words, detect a binary operator appearing at the beginning of an expression. Report that as an error, but also parse and discard a right-hand operand with the appropriate precedence.

Chapter 7 Evaluating Expressions

NOTS

1. In Lox, values are created by literals, computed by expressions, and stored in variables.

2. A literal is a bit of syntax that produces a value. A literal always appears somewhere in the user's source code. Lots of values are produced by computation and don't exist anywhere in the code itself. Those aren't literals. A literal comes from the parser's domain. Values are an interpreter concept, part of the runtime's world.

3. We can't evaluate the unary operator itself until after we evaluate its operand subexpression. That means our interpreter is doing a post-order traversal—each node evaluates its children before doing its own work.

4. Lox follows Ruby's simple rule: false and nil are falsey and everything else is truthy.

5. Runtime errors are failures that the language semantics demand we detect and report while the program is running (hence the name).

CHALLENGES

1. Allowing comparisons on types other than numbers could be useful. The operators might have a reasonable interpretation for strings. Even comparisons among mixed types, like 3 < "pancake" could be handy to enable things like ordered collections of heterogeneous types. Or it could

simply lead to bugs and confusion. Would you extend Lox to support comparing other types? If so, which pairs of types do you allow and how do you define their ordering? Justify your choices and compare them to other languages.

2. Many languages define + such that if either operand is a string, the other is converted to a string and the results are then concatenated. For example, "scone" + 4 would yield scone4. Extend the code in visitBinaryExpr() to support that.

edit visitBinaryExpr() function in Interpreter.java

```
1   // < Statements and State visit-assign
2     // > visit-binary
3     @Override
4     public Object visitBinaryExpr(Expr.Binary expr) {
5       Object left = evaluate(expr.left);
6       Object right = evaluate(expr.right); // [left]
7
8       switch (expr.operator.type) {
9         case GREATER:
10          // checkNumberOperands(expr.operator, left, right);
11          // accept compare number with the length of string
12          if (left instanceof Double && right instanceof Double)
13            return (double) left > (double) right;
14          if (left instanceof Double && right instanceof String)
15            return (double) left > right.toString().length();
16          if (left instanceof String && right instanceof Double)
17            return left.toString().length() > (double) right;
18          break;
19        case GREATER_EQUAL:
20          // checkNumberOperands(expr.operator, left, right);
21          // accept compare number with the length of string
22          if (left instanceof Double && right instanceof Double)
23            return (double) left >= (double) right;
24          if (left instanceof Double && right instanceof String)
25            return (double) left >= right.toString().length();
26          if (left instanceof String && right instanceof Double)
27            return left.toString().length() >= (double) right;
28          break;
29        case LESS:
30          // checkNumberOperands(expr.operator, left, right);
31          // accept compare number with the length of string
32          if (left instanceof Double && right instanceof Double)
33            return (double) left < (double) right;
34          if (left instanceof Double && right instanceof String)
35            return (double) left < right.toString().length();
36          if (left instanceof String && right instanceof Double)
37            return left.toString().length() < (double) right;
38          break;
39        case LESS_EQUAL:
```

```java
40        // checkNumberOperands(expr.operator, left, right);
41        // accept compare number with the length of string
42      if (left instanceof Double && right instanceof Double)
43        return (double) left <= (double) right;
44      if (left instanceof Double && right instanceof String)
45        return (double) left <= right.toString().length();
46      if (left instanceof String && right instanceof Double)
47        return left.toString().length() <= (double) right;
48      break;
49    case BANG_EQUAL:
50      return !isEqual(left, right);
51    case EQUAL_EQUAL:
52      return isEqual(left, right);
53    case MINUS:
54      checkNumberOperands(expr.operator, left, right);
55      return (double) left - (double) right;
56    case PLUS:
57      if (left instanceof Double && right instanceof Double) {
58        return (double) left + (double) right;
59      }
60      if (left instanceof String && right instanceof String) {
61        return (String) left + (String) right;
62      }
63      // make the lox accept add strings with numbers
64      if (left instanceof Double && right instanceof String) {
65        return stringify((Double) left) + (String) right;
66      }
67      if (left instanceof String && right instanceof Double) {
68        return (String) left + stringify((Double) right);
69      }
70      break;
71    // throw new RuntimeError(expr.operator,
72    // "Operands must be two numbers or two strings.");
73    case SLASH:
74      checkNumberOperands(expr.operator, left, right);
75      // handel our own messages on not allowed divisions
76      Object result = (double) left / (double) right;
77      if (result.toString() == "Infinity")
78        return "Division by zero is not allowed";
79      if (result.toString() == "NaN")
80        return "Not a Number";
81      return result;
82    case STAR:
83      // checkNumberOperands(expr.operator, left, right);
84      // return (double) left * (double) right;
85      if (left instanceof Double && right instanceof Double) {
86        return (double) left * (double) right;
87      }
88
89      if (left instanceof Double && right instanceof String) {
90        double repeatCount = (Double) left;
```

```
 91             if (repeatCount < 0) {
 92               throw new RuntimeError(null,
 93                   "Cannot repeat a string a negative number of times");
 94             }
 95             StringBuilder starResult = new StringBuilder();
 96             for (int i = 0; i < repeatCount; i++) {
 97               starResult.append(right);
 98             }
 99             return starResult.toString();
100           }
101       }
102     }
```

3. What happens right now if you divide a number by zero? What do you think should happen? Justify your choice. How do other languages you know handle division by zero and why do they make the choices they do? Change the implementation in visitBinaryExpr() to detect and report a runtime error for this case.

> like java and we can change this message in stringify() function

```
1  print 5 /0; // Infinity
2  print 0/0; // NaN
3  print 0/5; // 0
```