Atitit.跨语言scan parse ast api兼容性与标准化

Opcode是一种PHP脚本编译后的中间语言，就像Java的ByteCode,或者.NET的MSL，举个例子，比如你写下了如下的PHP代码：

1. <?php
2. echo "Hello World";
3. $a = 1 + 1;
4. echo $a;
5. ?>

PHP执行这段代码会经过如下4个步骤(确切的来说，应该是PHP的语言引擎Zend)

1. 1.Scanning(Lexing) ,将PHP代码转换为语言片段(Tokens)
2. 2.Parsing, 将Tokens转换成简单而有意义的表达式
3. 3.Compilation, 将表达式编译成Opocdes
4. 4.Execution, 顺次执行Opcodes，每次一条，从而实现PHP脚本的功能。

HP4.2开始提供了一个函数叫token\_get\_all,这个函数就可以讲一段PHP代码 Scanning成Tokens；  
如果用这个函数处理我们开头提到的PHP代码，将会得到如下结果:

如果用这个函数处理我们开头提到的PHP代码，将会得到如下结果:

1. Array
2. (
3. [0] => Array
4. (
5. [0] => 367
6. [1] => Array
7. (
8. [0] => 316
9. [1] => echo
10. )
11. [2] => Array
12. (
13. [0] => 370
14. [1] =>
15. )
16. [3] => Array
17. (
18. [0] => 315
19. [1] => "Hello World"
20. )
21. [4] => ;
22. [5] => Array
23. (
24. [0] => 370
25. [1] =>
26. )
27. [6] => =
28. [7] => Array
29. (
30. [0] => 370
31. [1] =>
32. )
33. [8] => Array
34. (
35. [0] => 305
36. [1] => 1
37. )
38. [9] => Array
39. (
40. [0] => 370
41. [1] =>
42. )
43. [10] => +
44. [11] => Array
45. (
46. [0] => 370
47. [1] =>
48. )
49. [12] => Array
50. (
51. [0] => 305
52. [1] => 1
53. )
54. [13] => ;
55. [14] => Array
56. (
57. [0] => 370
58. [1] =>
59. )
60. [15] => Array
61. (
62. [0] => 316
63. [1] => echo
64. )
65. [16] => Array
66. (
67. [0] => 370
68. [1] =>
69. )
70. [17] => ;
71. )

分析这个返回结果我们可以发现，源码中的字符串，字符，空格，都会原样返回。每个源代码中的字符，都会出现在相应的顺序处。而，其他的比如标签，操作符，语句，都会被转换成一个包含俩部分的Array: Token ID (也就是在Zend内部的改Token的对应码，比如,T\_ECHO,T\_STRING)，和源码中的原来的内容。  
接下来，就是Parsing阶段了，Parsing首先会丢弃Tokens Array中的多于的空格，然后将剩余的Tokens转换成一个一个的简单的表达式

1. 1.echo a constant string
2. 2.add two numbers together
3. 3.store the result of the prior expression to a variable
4. 4.echo a variable

然后就改Compilation阶段了，它会把Tokens编译成一个个op\_array, 每个op\_arrayd包含如下5个部分：

1. 1.Opcode数字的标识，指明了每个op\_array的操作类型，比如add , echo
2. 2.结果 存放Opcode结果
3. 3.操作数1 给Opcode的操作数
4. 4.操作数2
5. 5.扩展值 1个整形用来区别被重载的操作符

比如，我们的PHP代码会被Parsing成:

1. \* ZEND\_ECHO 'Hello World'
2. \* ZEND\_ADD ~0 1 1
3. \* ZEND\_ASSIGN !0 ~0
4. \* ZEND\_ECHO !0

参考

深入理解PHP原理之Opcodes 风雪之隅.htm

php用vld查看opcode代码\_听风\_新浪博客.htm

PHP中的NOP及为什么有这个opcode -- 编程语言 -- IT技术博客大学习 -- 共学习 共进步！.htm

PHP VLD扩展查看opcode - PHP点点通.htm