Semantic Analysis

TEACHING ASSISTANT: DAVID TRABISH

Semantic Analysis

We need to check the following:

- Type checking
 - 1 + "1"
- Scopes
 - Undefined variables
- Other
 - Division by zero
 - Const variables
 - Visibility semantics in classes (public, private, ...)

Symbol Table

- Maintain a stack of scopes
- Each scope maps identifiers to their type information
- Identifiers may be:
 - Variable names
 - Function names
 - Method names

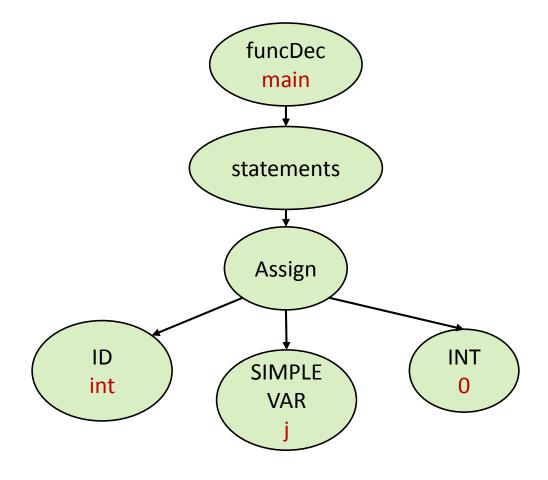
Symbol Table

- When we reach a variable/function/... declaration
 - Update the map of the current scope (top of the stack)
- When we reach a new block, **push** a new scope
- When we leave a block, **pop** the top scope
- Begin with the global (initial scope)
 - Functions, global variables, ...

Symbol Table

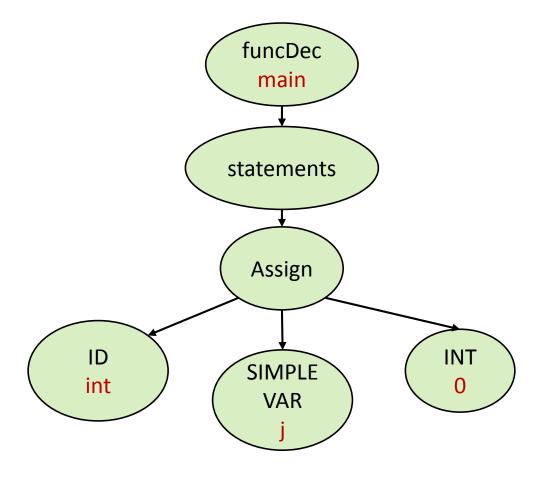
- When we need to resolve an identifier
 - Scan the scopes (starting from the top)
 - Stop at the first matching scope
 - If no scope was found, we have an error...

```
void main() {
  int j = 0;
}
```

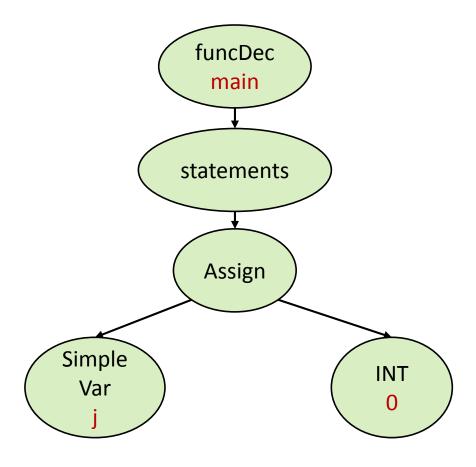


```
void main() {
  int j = 0;
}
```

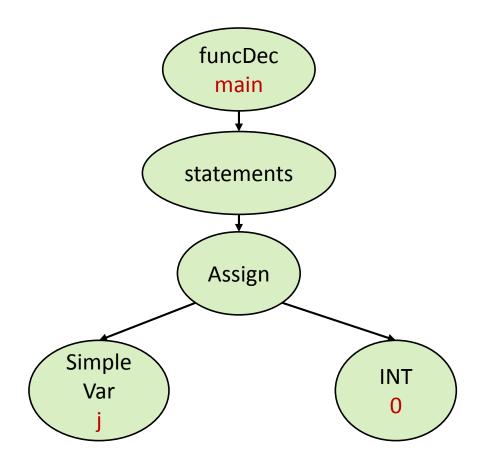
Valid



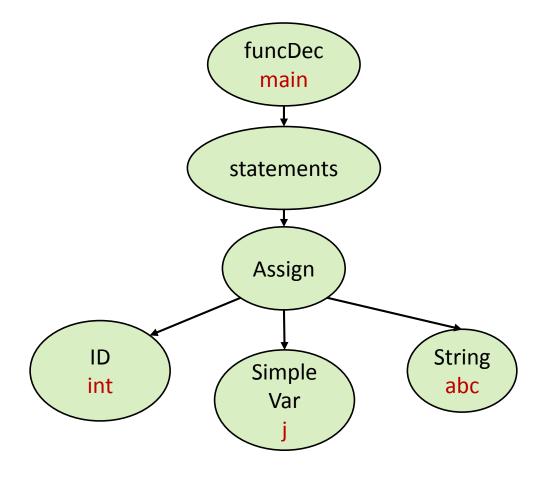
```
void main() {
   j = 0;
}
```



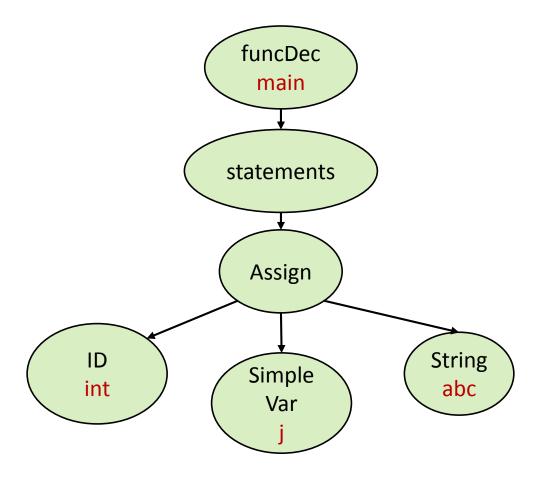
```
void main() {
   j = 0;
}
```



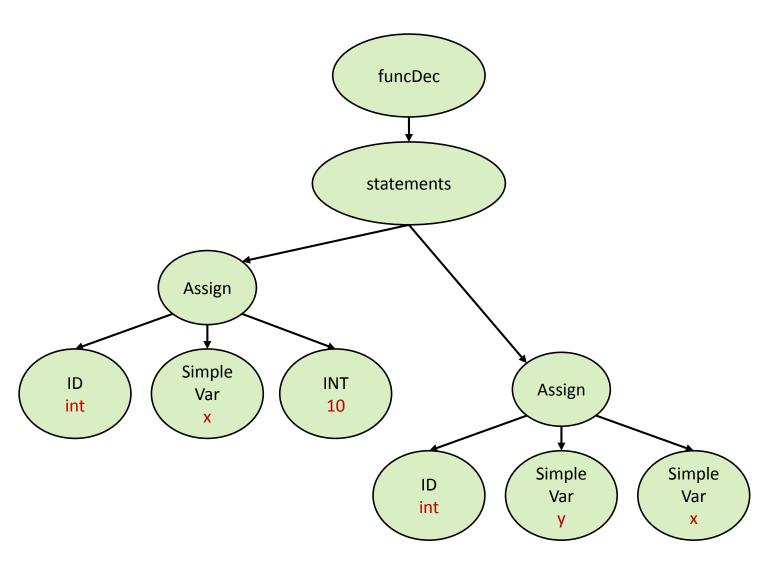
```
void main() {
  int j = "abc";
}
```



```
void main() {
  int j = "abc";
}
```

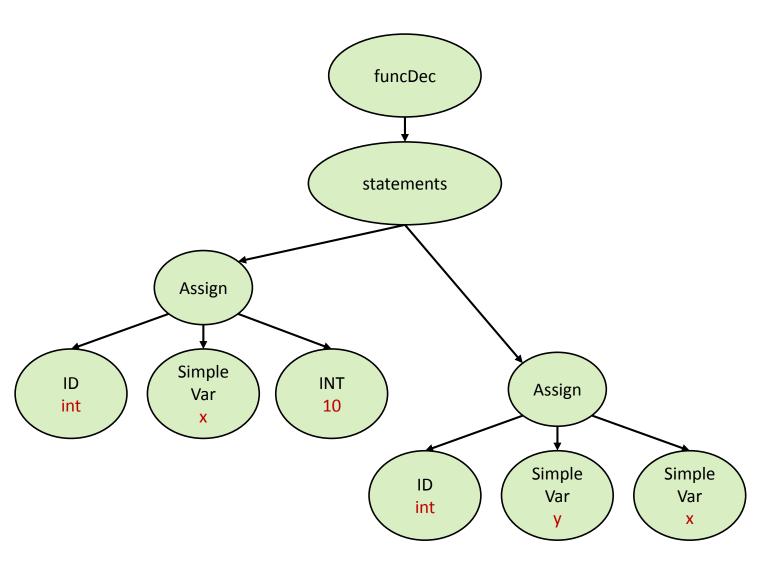


```
void main() {
  int x = 10;
  int y = x;
}
```

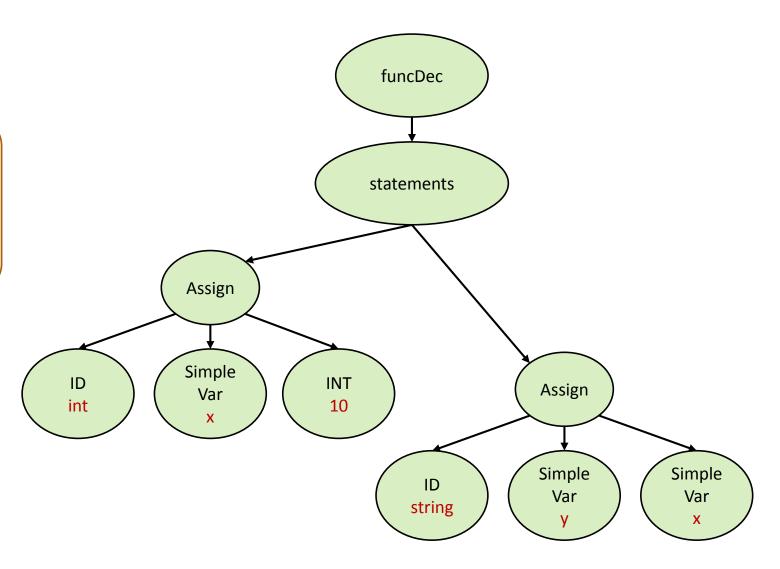


```
void main() {
  int x = 10;
  int y = x;
}
```

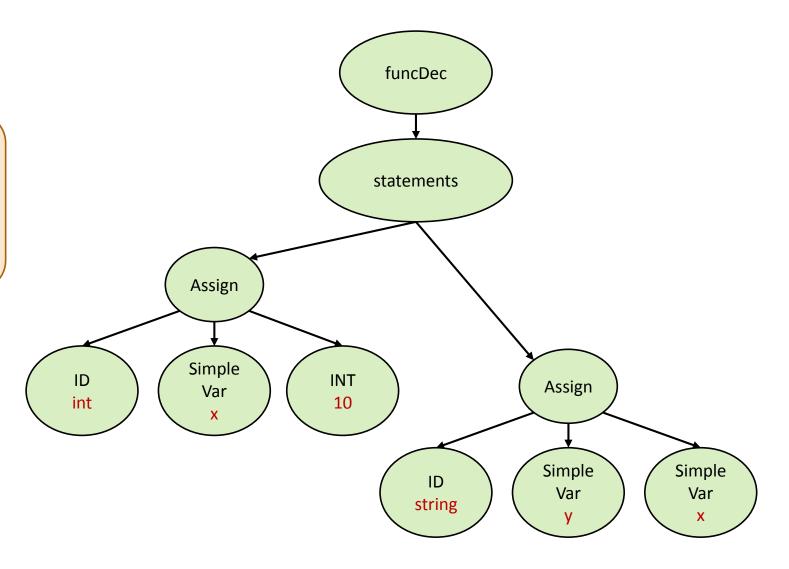
Valid



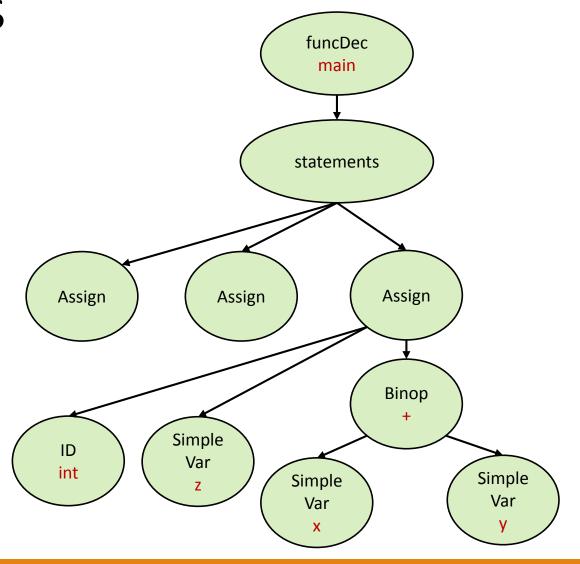
```
void main() {
  int x = 10;
  string y = x;
}
```



```
void main() {
  int x = 10;
  string y = x;
}
```

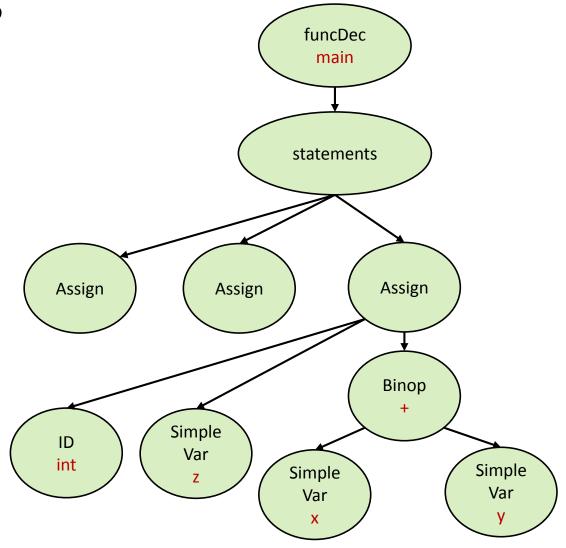


```
void main() {
  int x = 1;
  int y = 2;
  int z = x + y;
}
```

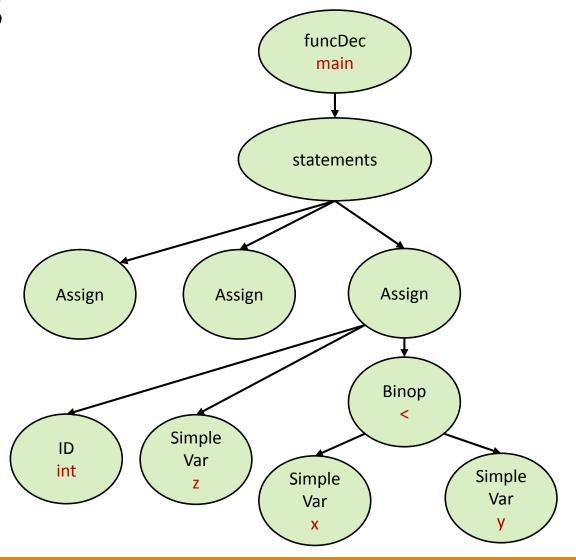


```
void main() {
  int x = 1;
  int y = 2;
  int z = x + y;
}
```

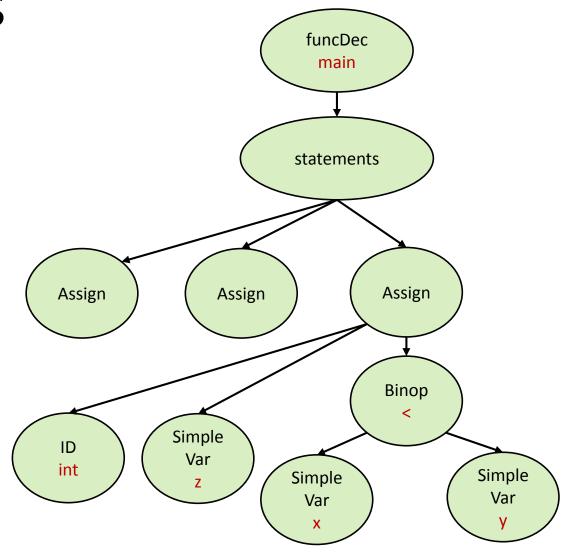
Valid



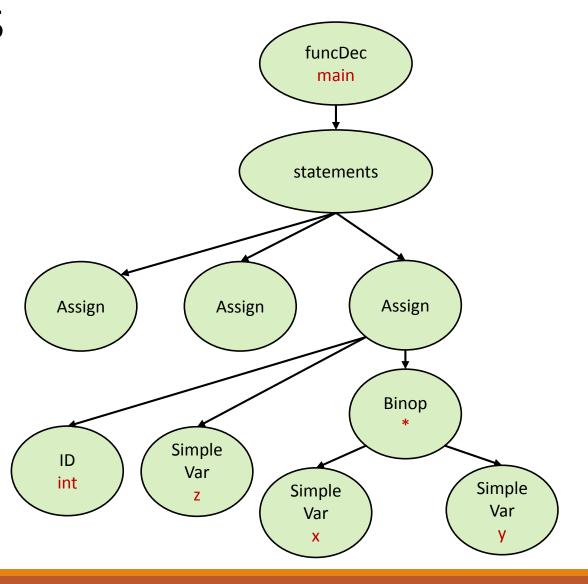
```
void main() {
  int x = 1;
  string y = "A";
  int z = x < y;
}</pre>
```



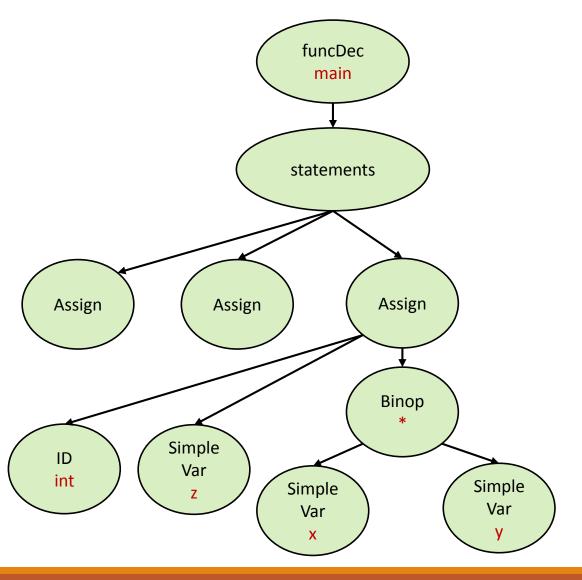
```
void main() {
  int x = 1;
  string y = "A";
  int z = x < y;
}</pre>
```



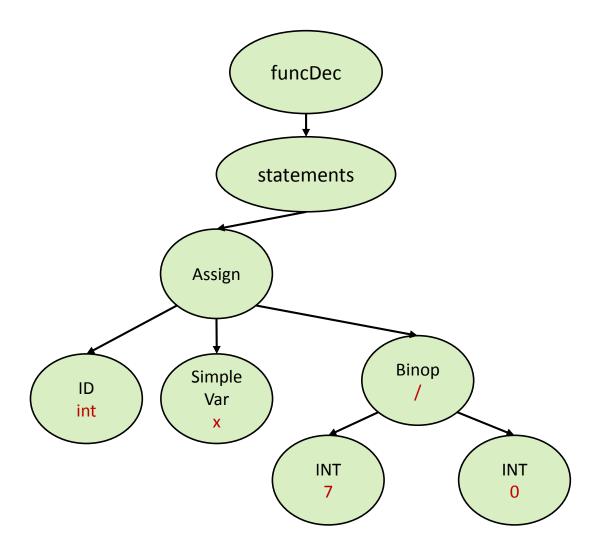
```
void main() {
  string x = "A";
  string y = "B";
  string z = x * y;
}
```



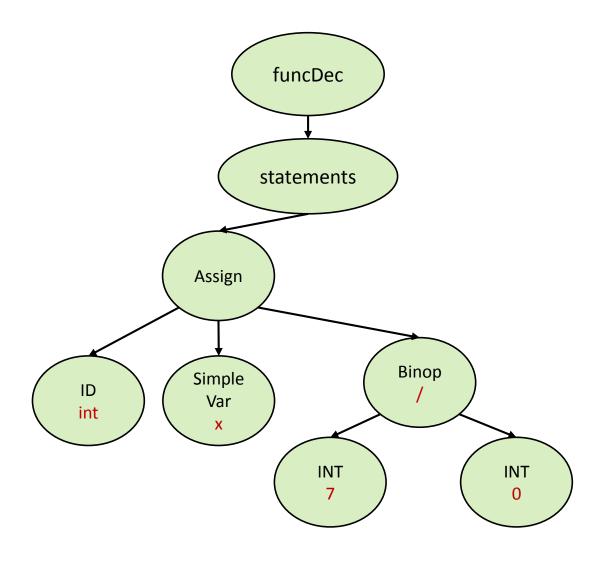
```
void main() {
  string x = "A";
  string y = "B";
  string z = x * y;
}
```



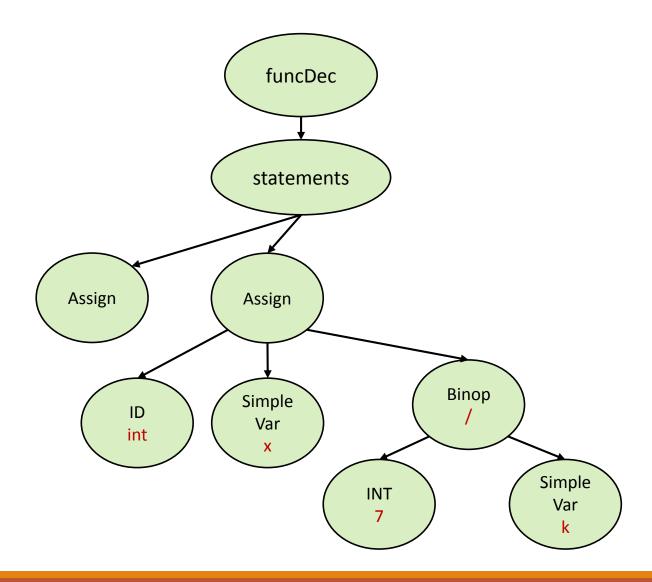
```
void main() {
  int x = 7 / 0;
}
```



```
void main() {
  int x = 7 / 0;
}
```

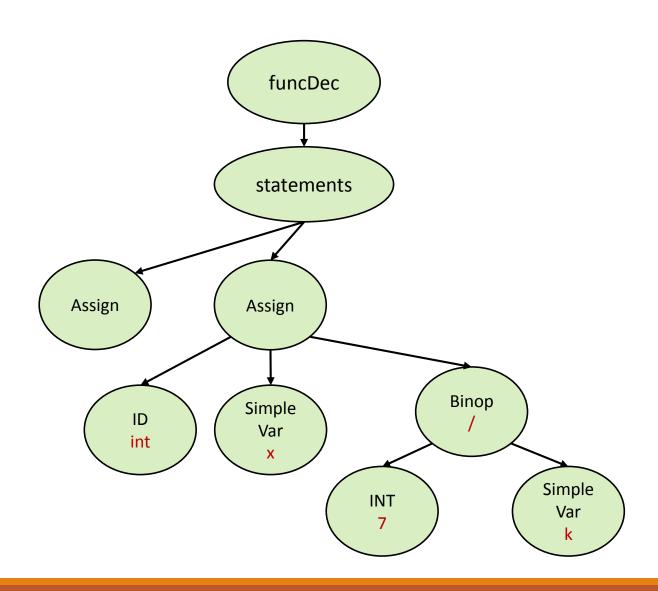


```
void main() {
  int k = 0;
  int x = 7 / k;
}
```

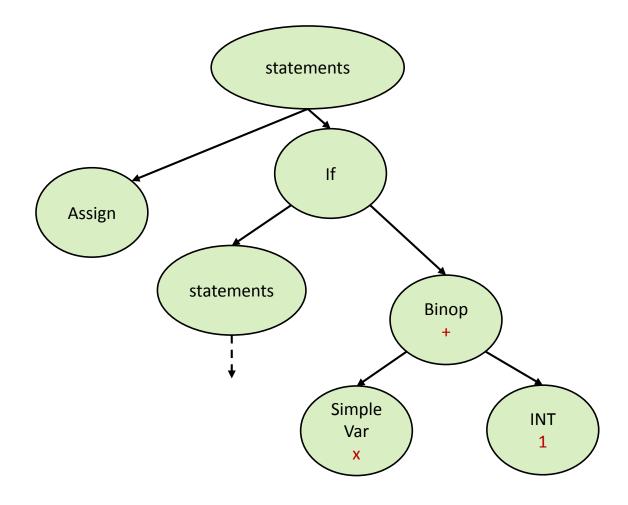


```
void main() {
  int k = 0;
  int x = 7 / k;
}
```

Depends

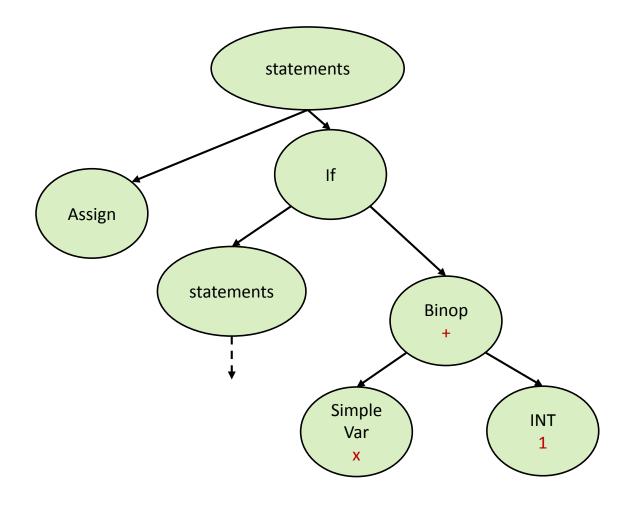


```
void main() {
  int x = 1;
  if (x + 1) {
    int z = 2;
  }
}
```

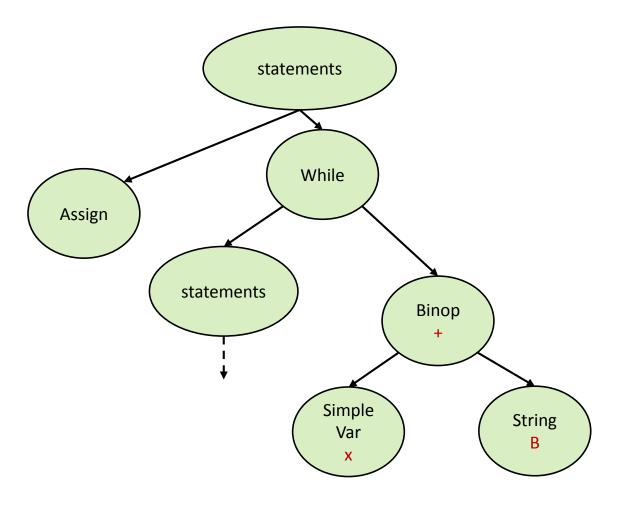


```
void main() {
  int x = 1;
  if (x + 1) {
    int z = 2;
  }
}
```

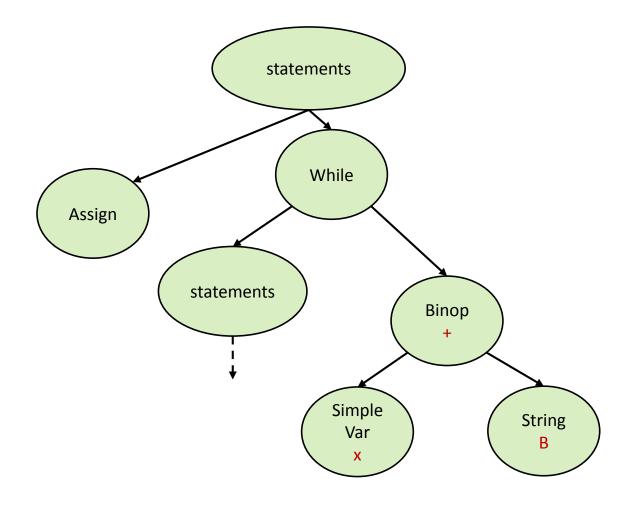
Valid



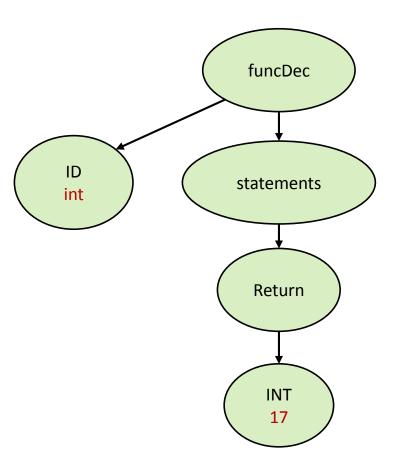
```
void main() {
   string x = "A";
   while (x + "B") {
     int z = 2;
   }
}
```



```
void main() {
   string x = "A";
   while (x + "B") {
     int z = 2;
   }
}
```

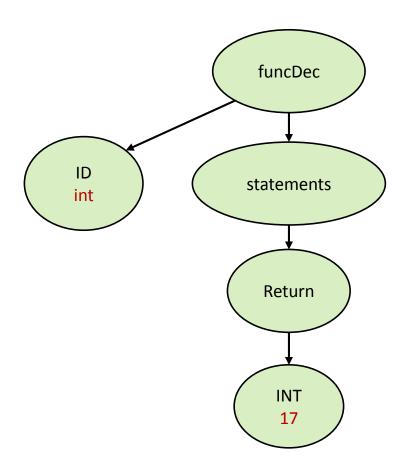


```
int main() {
  return 17;
}
```

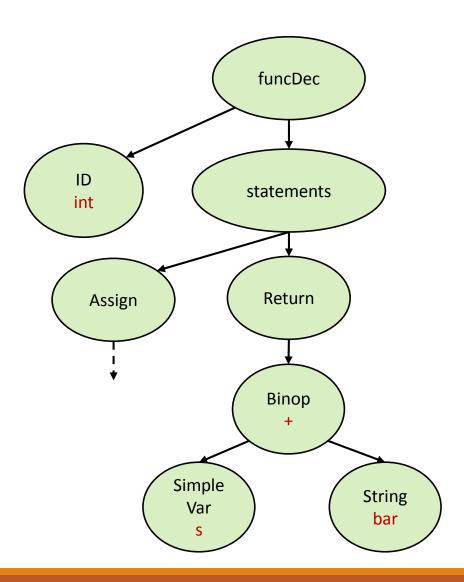


```
int main() {
  return 17;
}
```

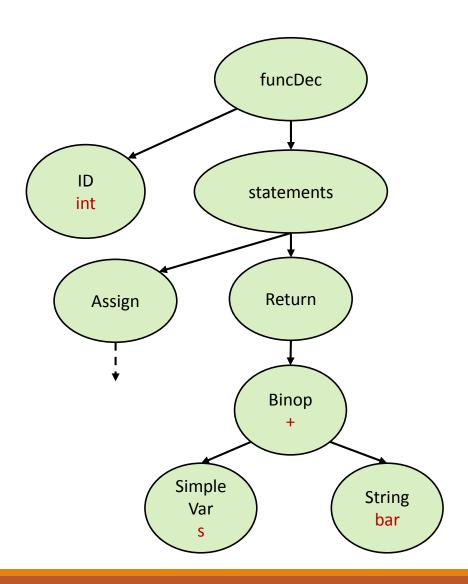
Valid



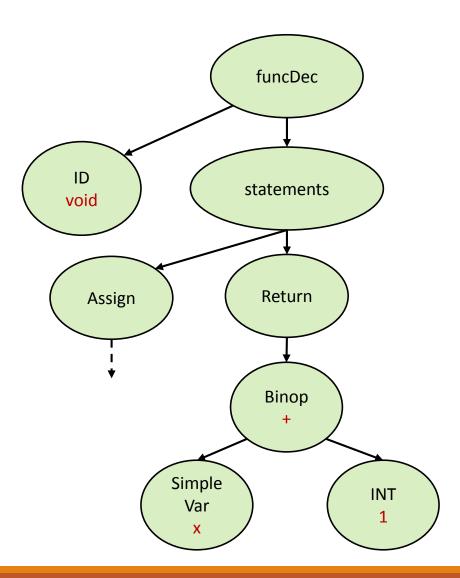
```
int main() {
   string s = "foo"
   return s + "bar";
}
```



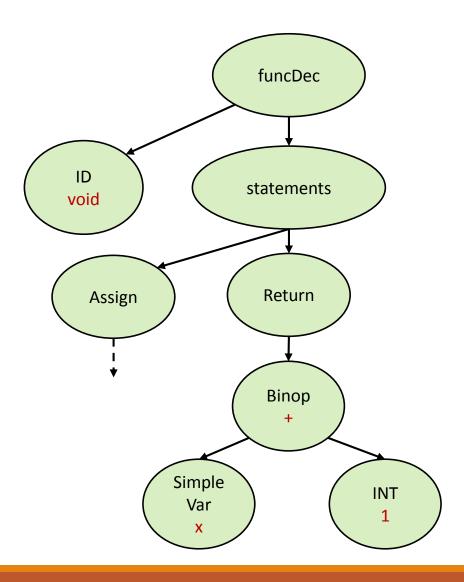
```
int main() {
   string s = "foo"
   return s + "bar";
}
```



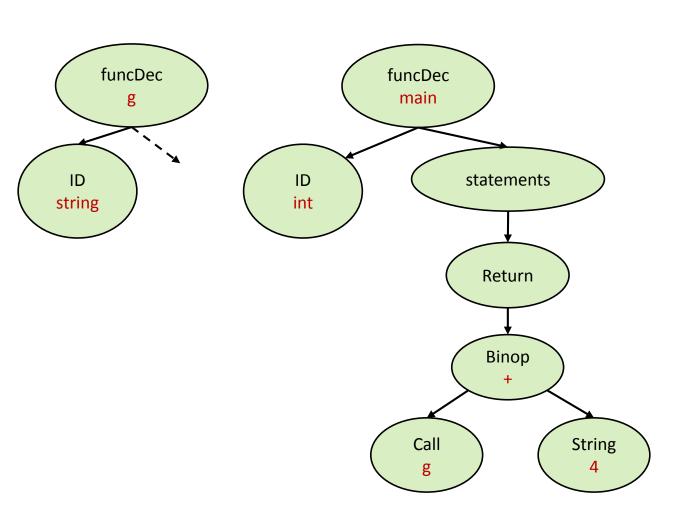
```
void main() {
  int x = 1;
  return x + 1;
}
```



```
void main() {
  int x = 1;
  return x + 1;
}
```



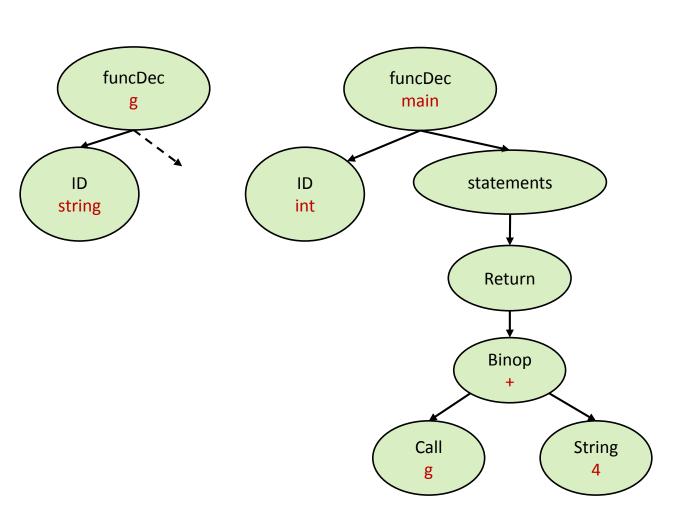
```
string g() {
  return "123";
}
int main() {
  return g() + "4";
}
```



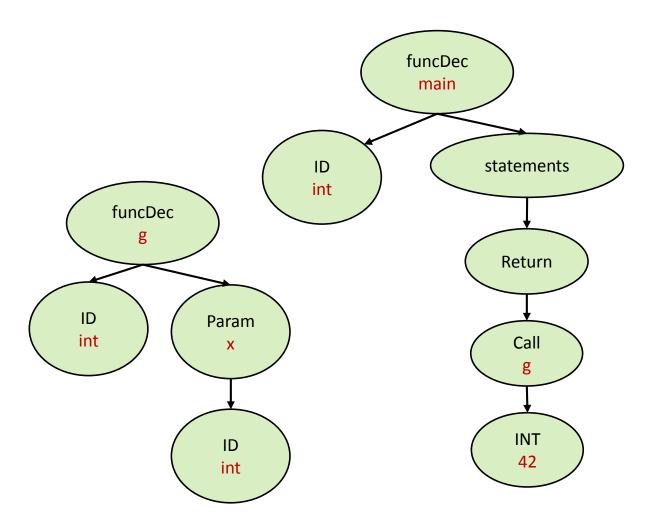
Return Statement

```
string g() {
  return "123";
}
int main() {
  return g() + "4";
}
```



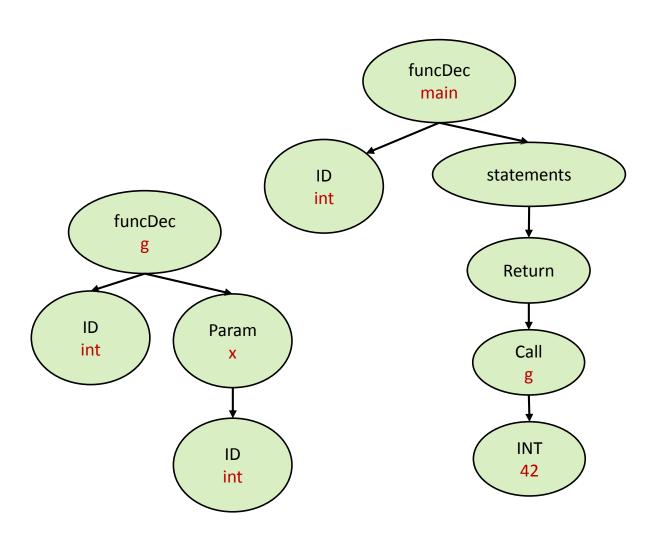


```
int g(int x) {
   return x + 1;
}
int main() {
   return g(42);
}
```

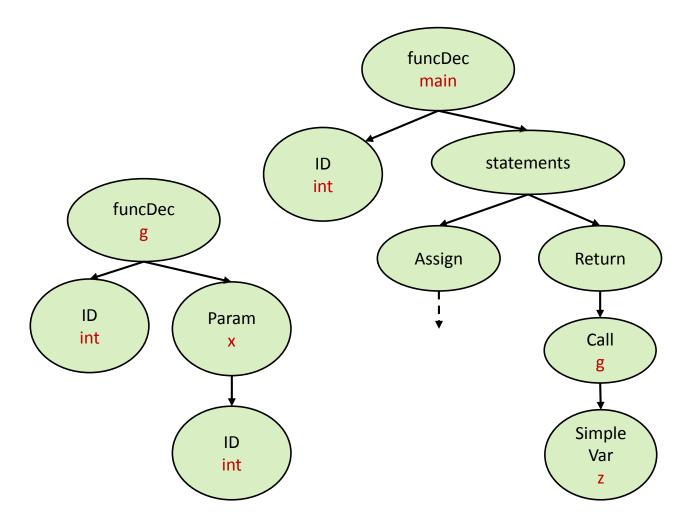


```
int g(int x) {
   return x + 1;
}
int main() {
   return g(42);
}
```

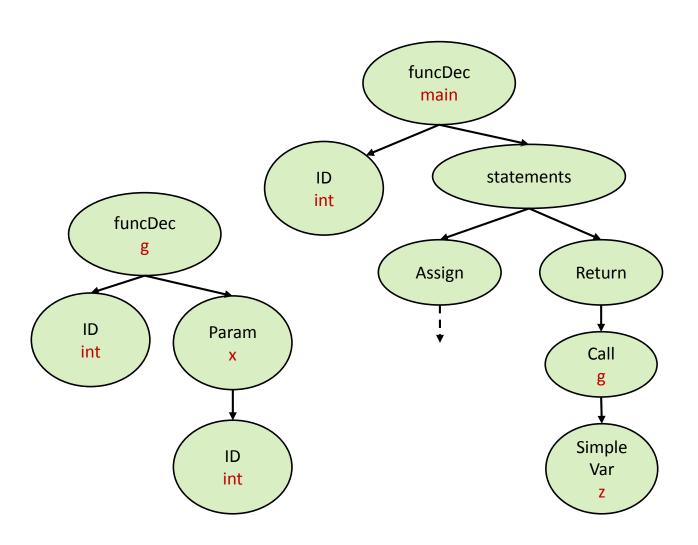
Valid



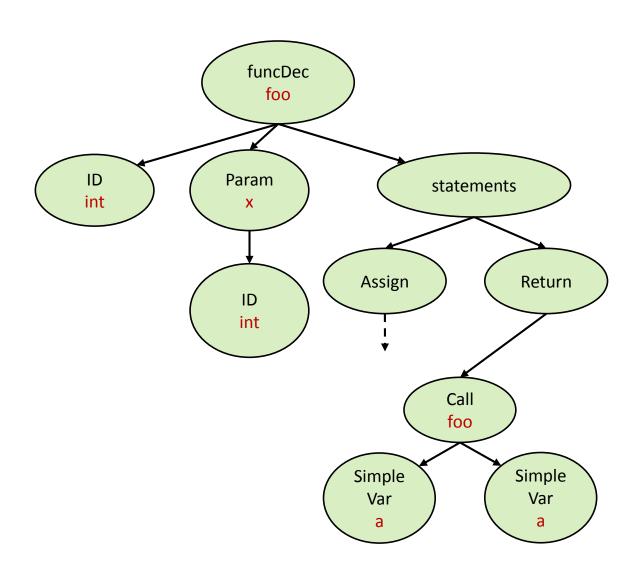
```
int g(int x) {
  return x + 1;
}
int main() {
  string z = "..."
  return g(z);
}
```



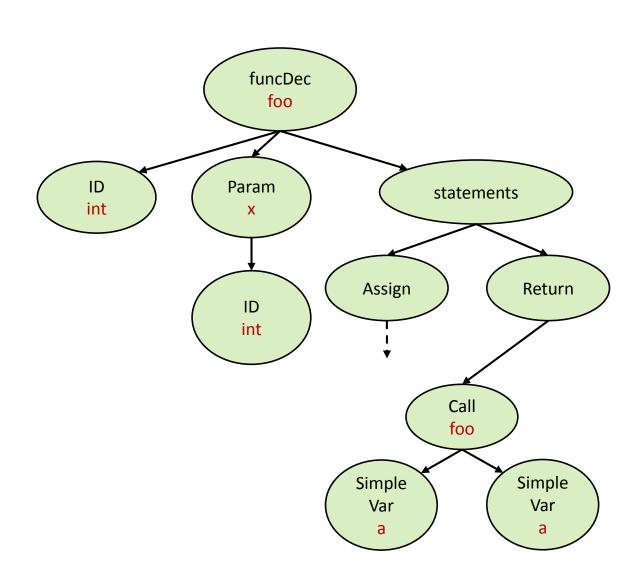
```
int g(int x) {
  return x + 1;
}
int main() {
  string z = "..."
  return g(z);
}
```



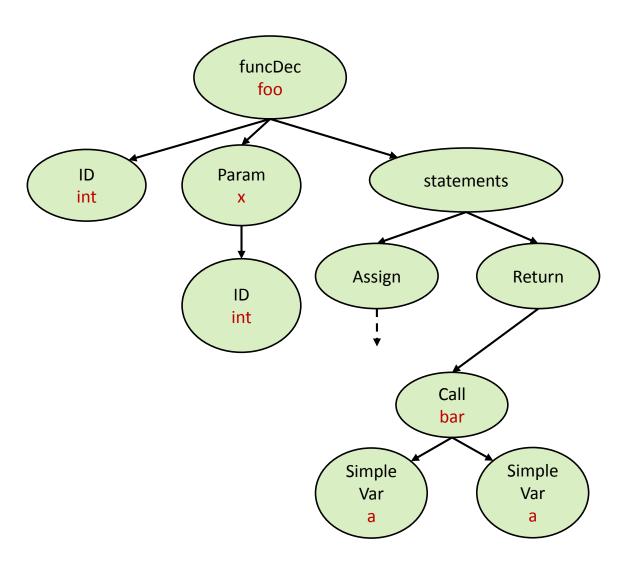
```
int foo(int k) {
  int a = k * 10;
  return foo(a, a);
}
```



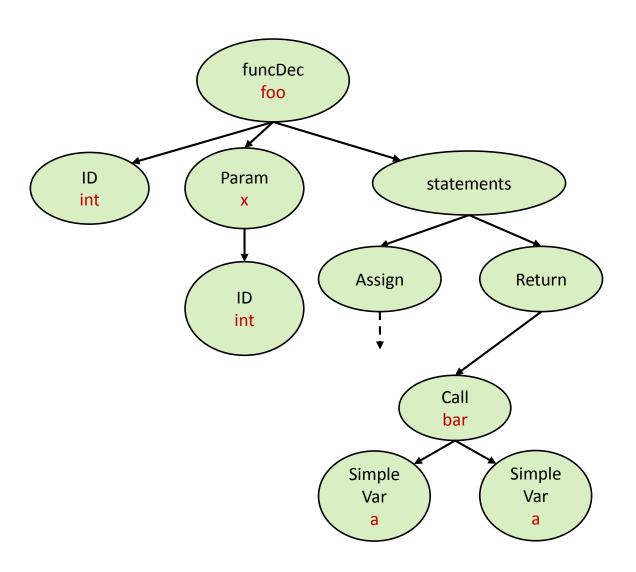
```
int foo(int k) {
  int a = k * 10;
  return foo(a, a);
}
```

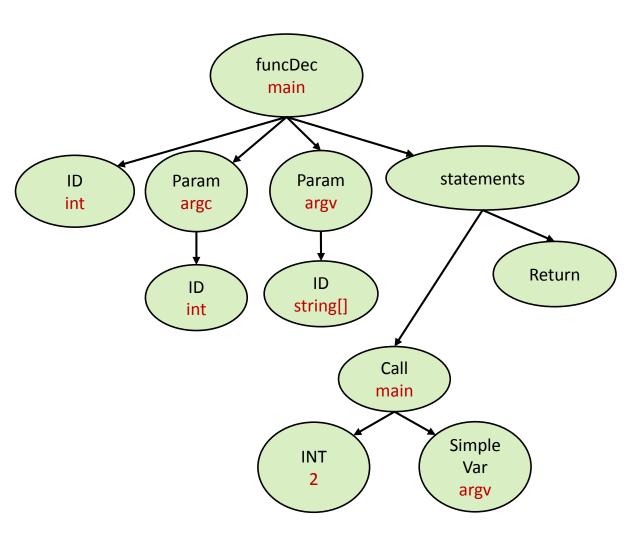


```
int foo(int k) {
  int a = k * 10;
  return bar(a, a);
}
```

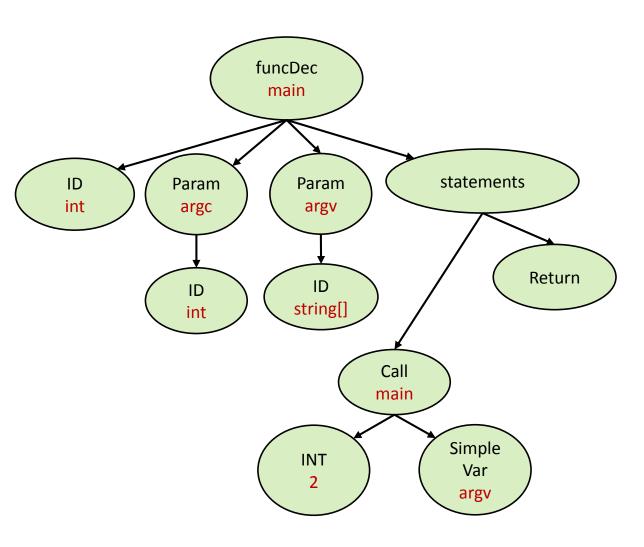


```
int foo(int k) {
  int a = k * 10;
  return bar(a, a);
}
```





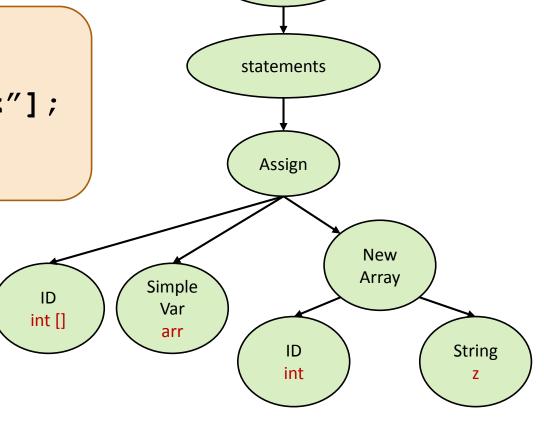
Valid



```
funcDec
                                                           foo
void foo(void) {
                                                         statements
   int[] arr = new int["z"];
                                                          Assign
                                                                   New
                                                                  Array
                                                 Simple
                                                  Var
                                       int []
                                                  arr
                                                                          String
```

```
void foo(void) {
  int[] arr = new int["z"];
}
```

Invalid

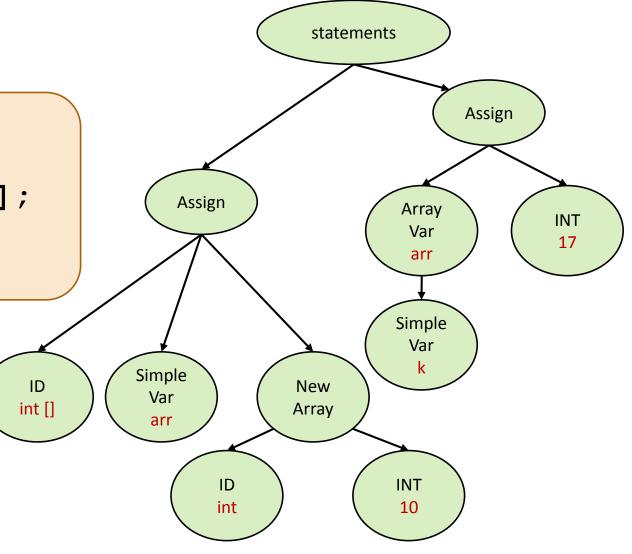


funcDec foo

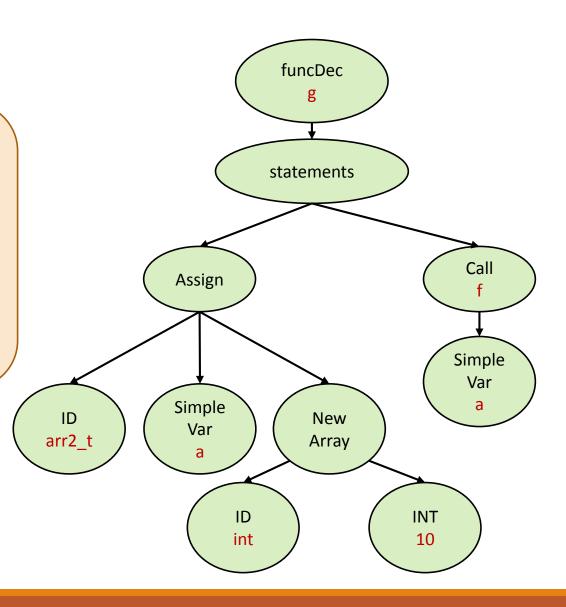
```
statements
void foo(int d) {
                                                                            Assign
   int k = 3;
   int[] arr = new int[10];
                                                     Assign
                                                                       Array
                                                                                    INT
   arr[k] = 17;
                                                                        Var
                                                                                    17
                                                                        arr
                                                                       Simple
                                                                        Var
                                                  Simple
                                        ID
                                                               New
                                                  Var
                                        int []
                                                              Array
                                                   arr
                                                                       INT
                                                        ID
                                                                       10
                                                        int
```

```
void foo(int d) {
  int k = 3;
  int[] arr = new int[10];
  arr[k] = 17;
}
```

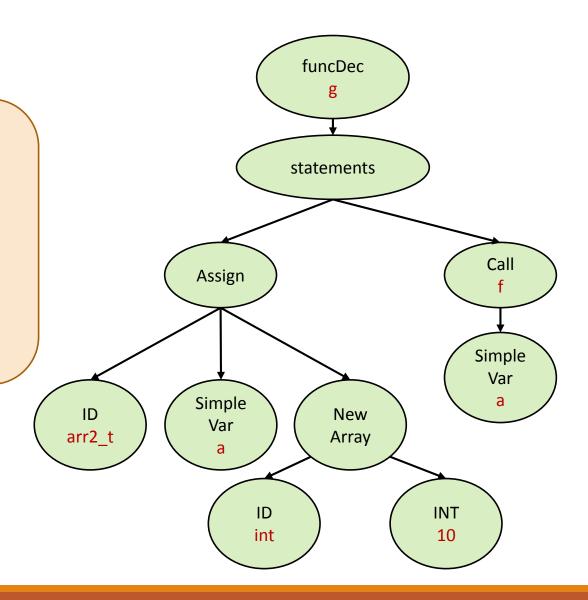
Valid



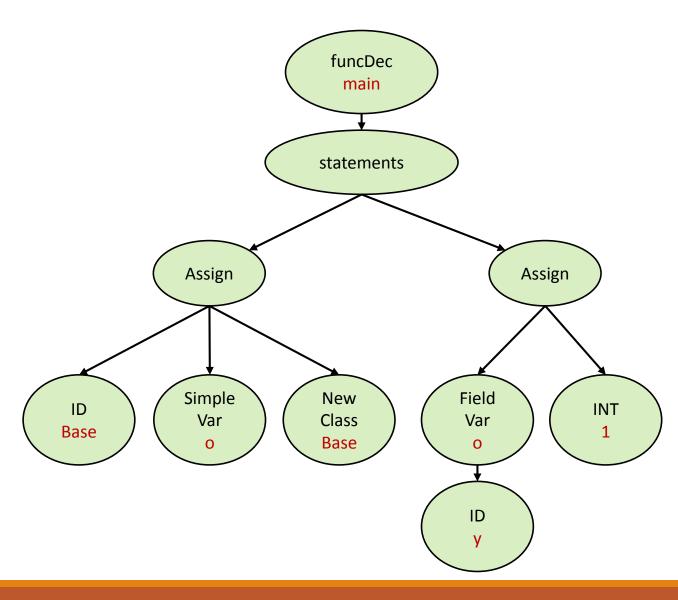
```
typedef int arr1_t[];
typedef int arr2_t[];
void f(arr1_t a) { }
void g() {
   arr2_t a = new int[10];
   f(a);
}
```



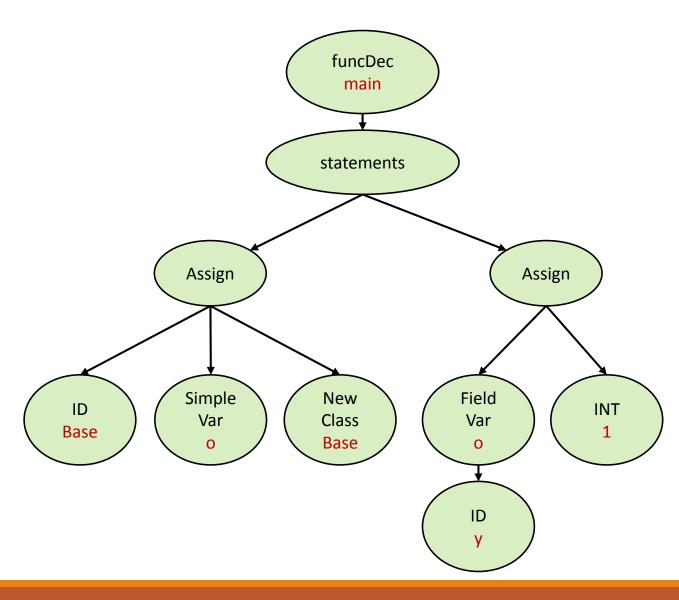
```
typedef int arr1_t[];
typedef int arr2_t[];
void f(arr1_t a) { }
void g() {
   arr2_t a = new int[10];
   f(a);
}
```



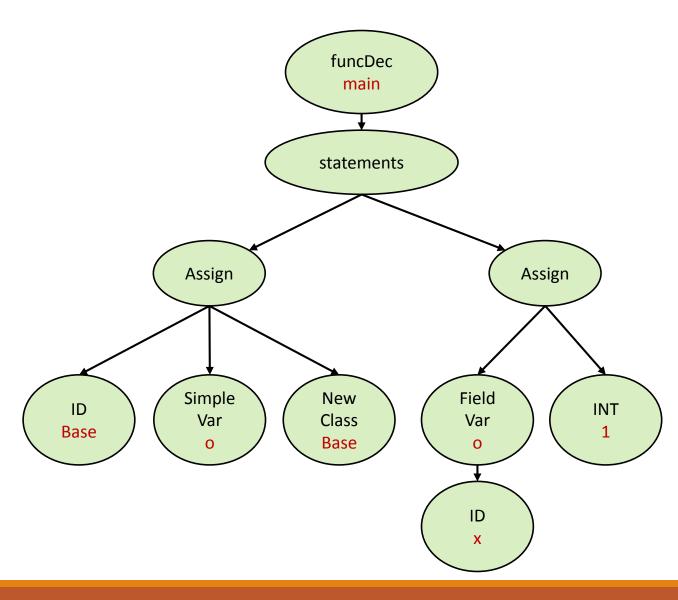
```
class Base {
  int x;
}
void main() {
  Base o = new Base;
  o.y = 1;
}
```



```
class Base {
  int x;
}
void main() {
  Base o = new Base;
  o.y = 1;
}
```

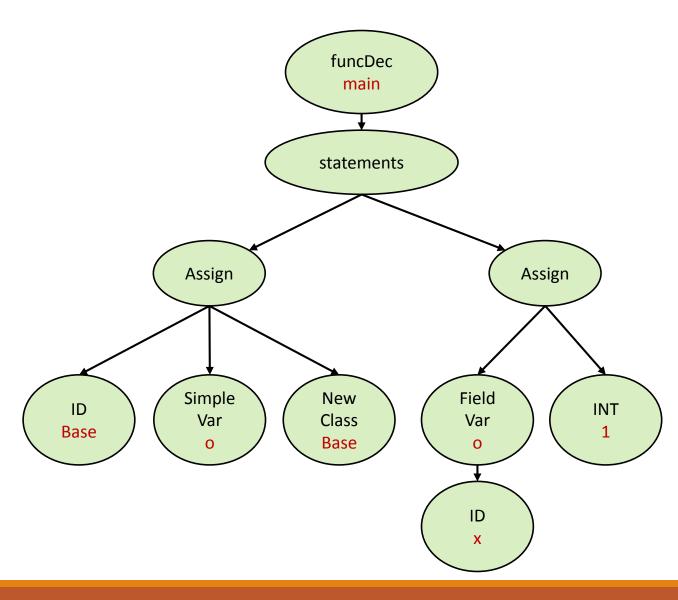


```
class Base {
  int x;
}
void main() {
  Base o = new Base;
  o.x = 1;
}
```

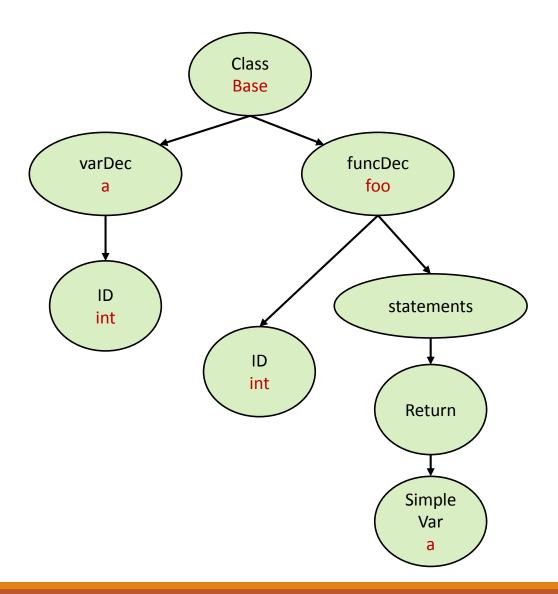


```
class Base {
  int x;
}
void main() {
  Base o = new Base;
  o.x = 1;
}
```

Valid

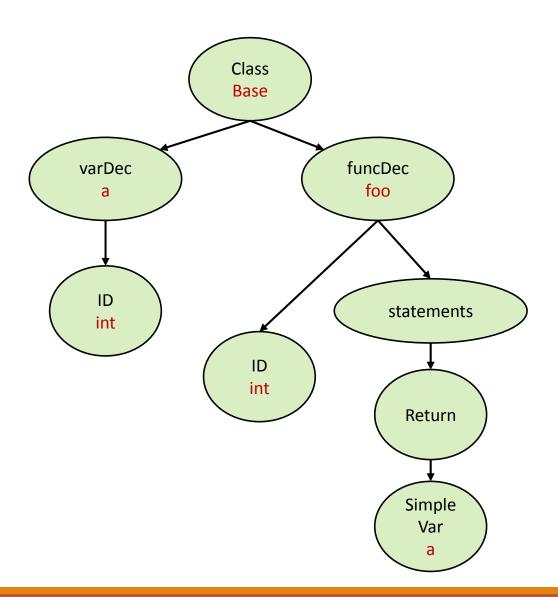


```
class Base {
  int a;
  int foo() {
    return a;
  }
}
```

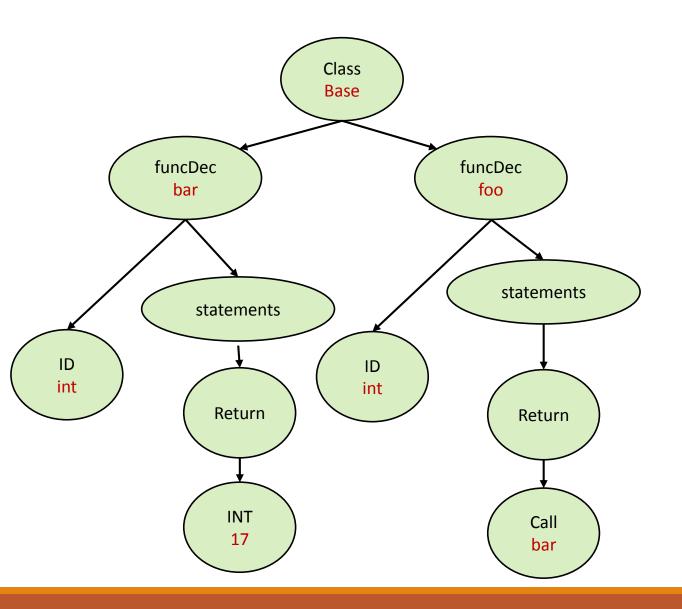


```
class Base {
  int a;
  int foo() {
    return a;
  }
}
```

Valid

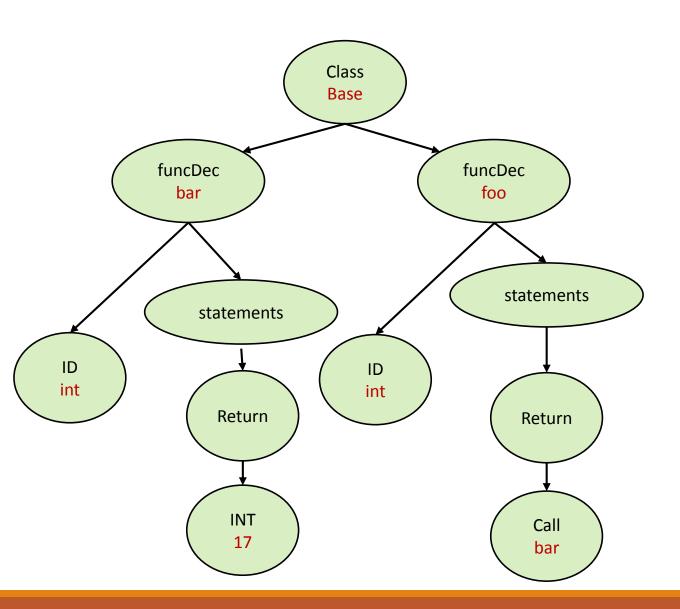


```
class Base {
  int bar() {
    return 17;
  }
  int foo() {
    return bar();
  }
}
```

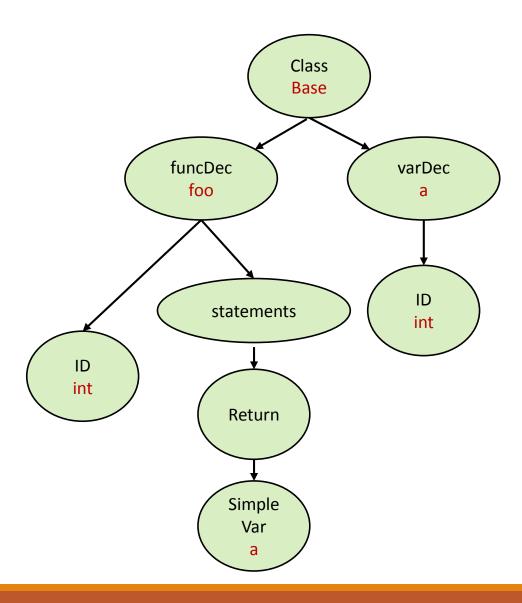


```
class Base {
  int bar() {
    return 17;
  }
  int foo() {
    return bar();
  }
}
```

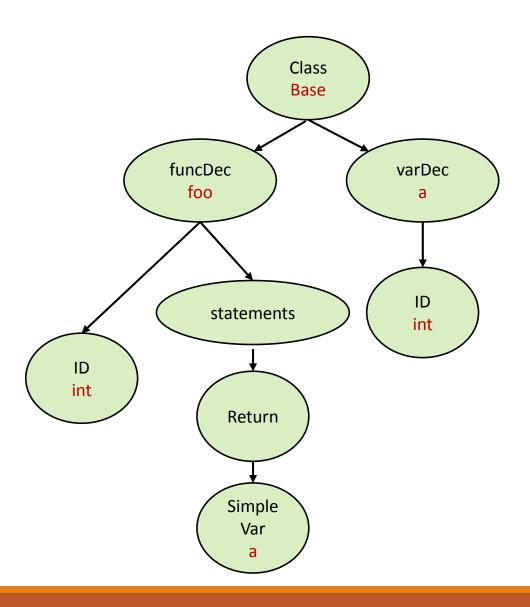
Valid



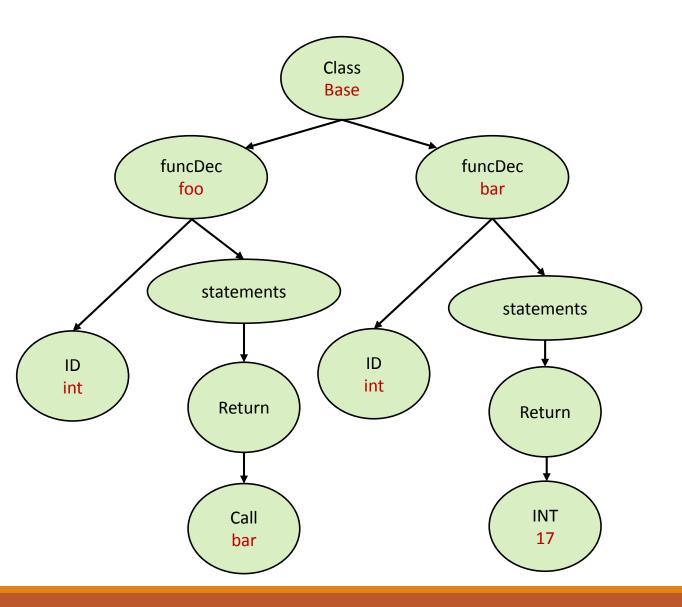
```
class Base {
  int foo() {
    return a;
  }
  int a;
}
```



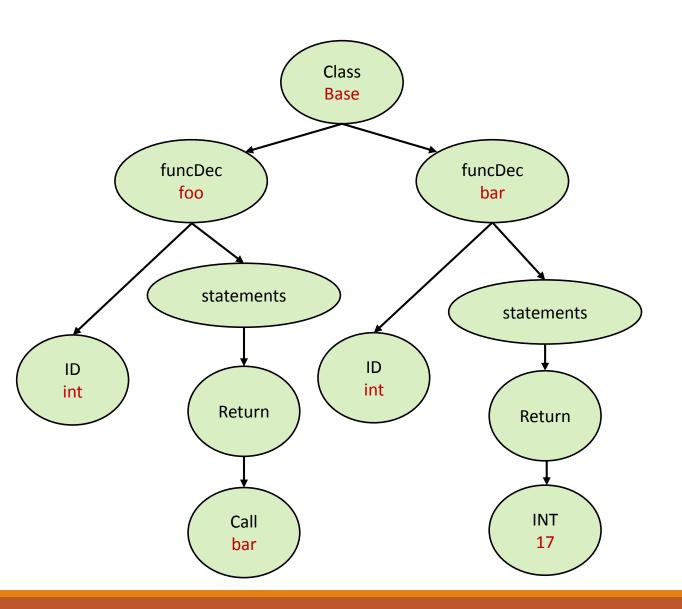
```
class Base {
  int foo() {
    return a;
  }
  int a;
}
```



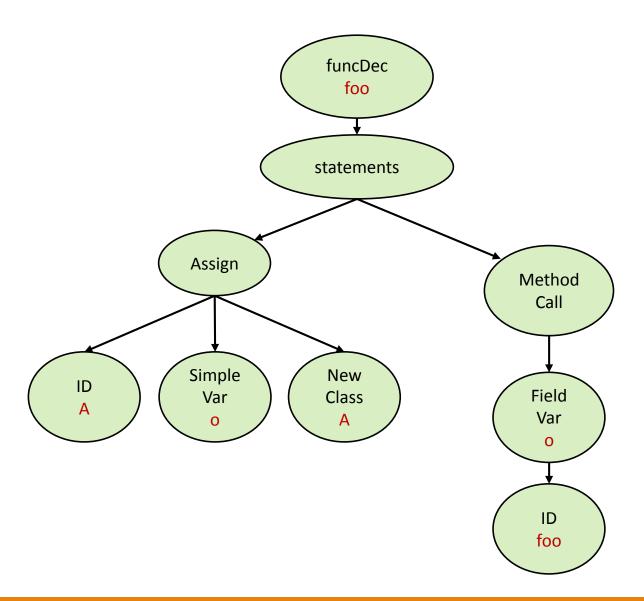
```
class Base {
  int foo() {
    return bar();
  }
  int bar() {
    return 17;
  }
}
```



```
class Base {
  int foo() {
    return bar();
  }
  int bar() {
    return 17;
  }
}
```

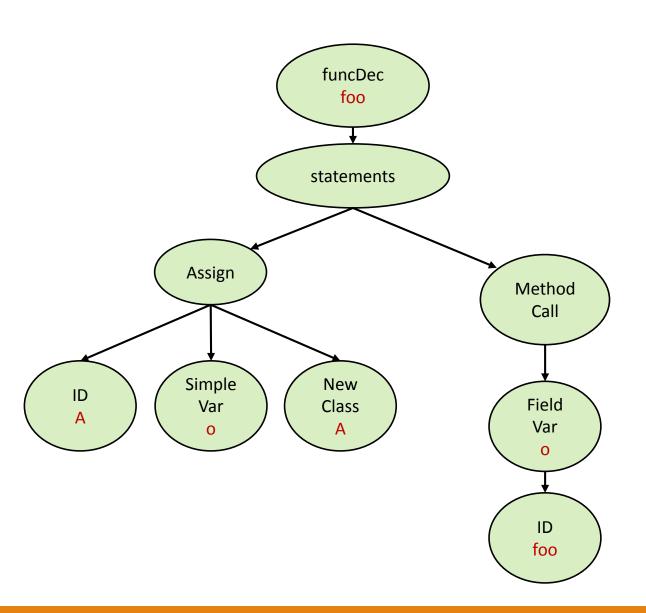


```
class A {
  void foo() {
    A o = new A;
    o.foo();
  }
}
```

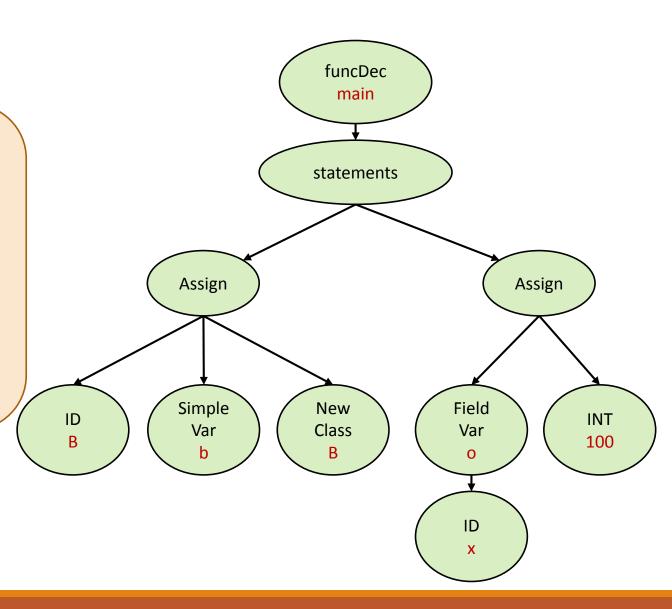


```
class A {
  void foo() {
    A o = new A;
    o.foo();
  }
}
```



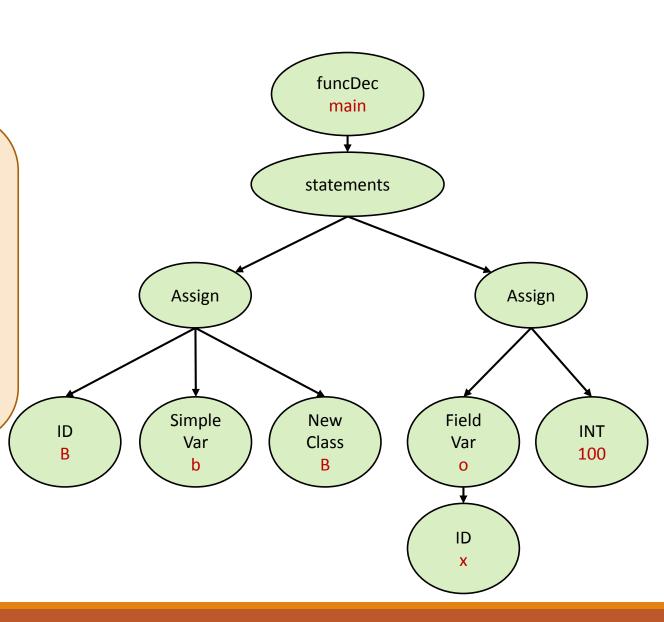


```
class A {
  int x;
}
class B extends A { }
void main() {
  B o = new B;
  o.x = 100;
}
```

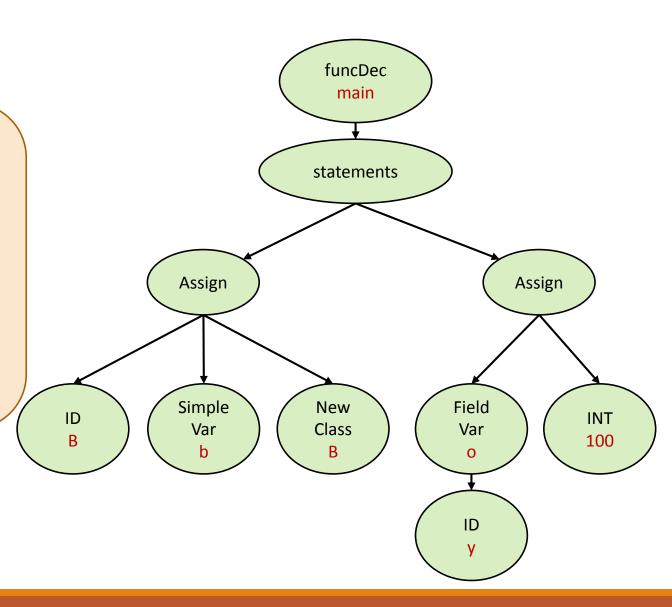


```
class A {
  int x;
}
class B extends A { }
void main() {
  B o = new B;
  o.x = 100;
}
```

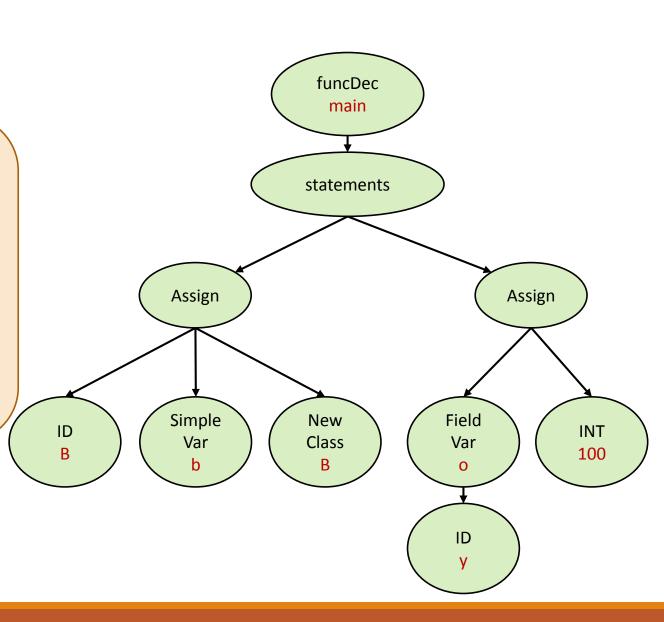




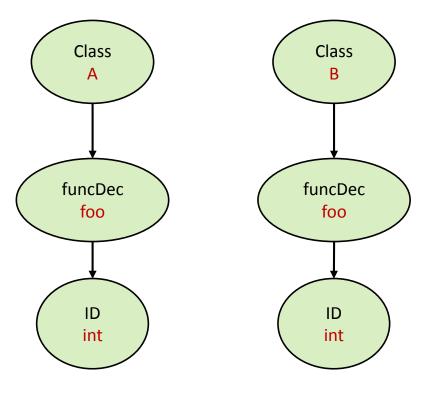
```
class A {
  int x;
}
class B extends A { }
void main() {
  B o = new B;
  o.y = 100;
}
```



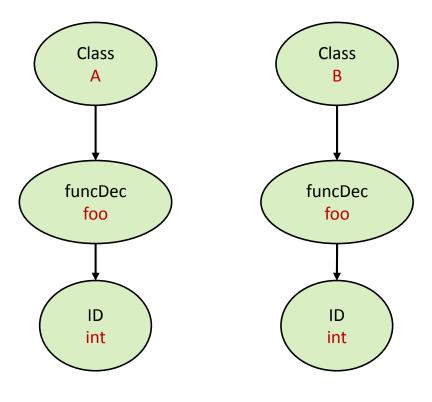
```
class A {
  int x;
}
class B extends A { }
void main() {
  B o = new B;
  o.y = 100;
}
```



```
class A {
  int foo() {
    return 17;
class B extends A {
  int foo() {
    return 18;
```

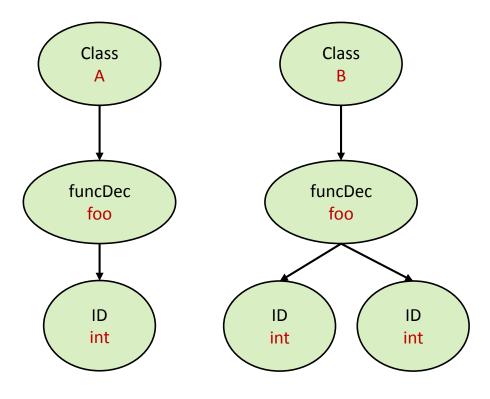


```
class A {
  int foo() {
    return 17;
class B extends A {
  int foo() {
    return 18;
```

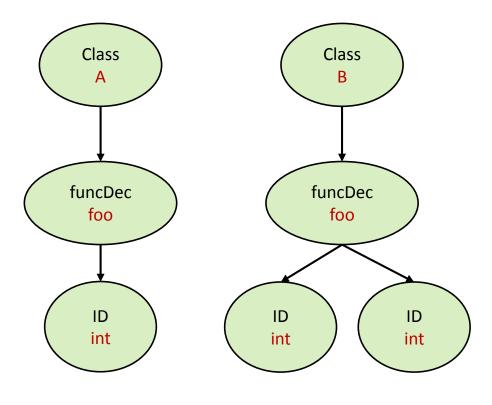




```
class A {
  int foo() {
    return 17;
class B extends A {
  int foo(int x) {
    return x + 1;
```

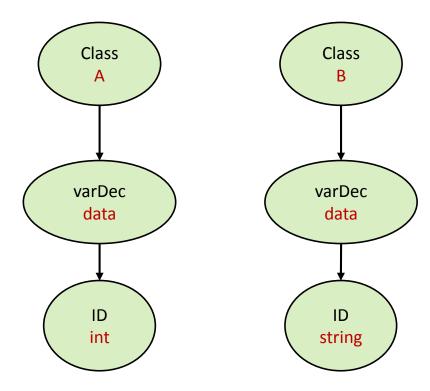


```
class A {
  int foo() {
    return 17;
class B extends A {
  int foo(int x) {
    return x + 1;
```

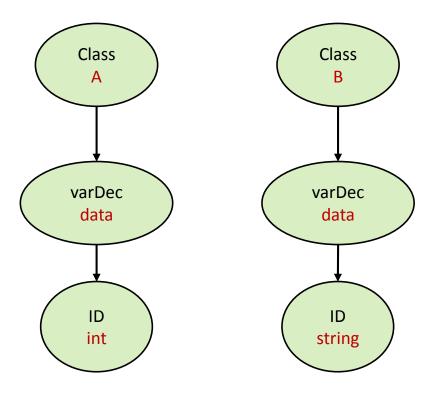


Invalid

```
class A {
  int data;
}
class B extends A {
  string data;
}
```

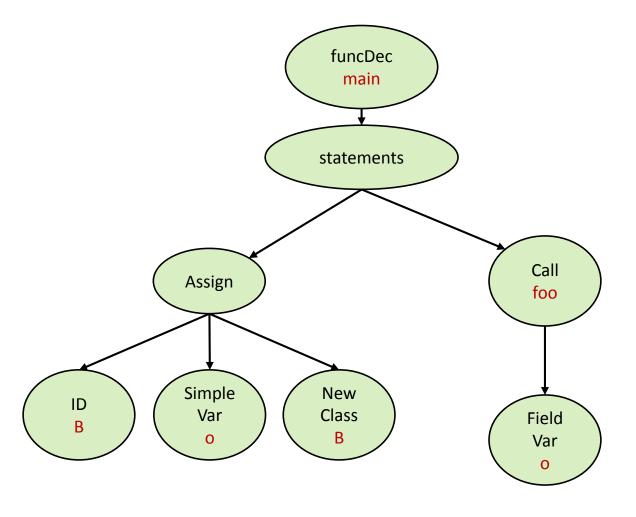


```
class A {
  int data;
}
class B extends A {
  string data;
}
```



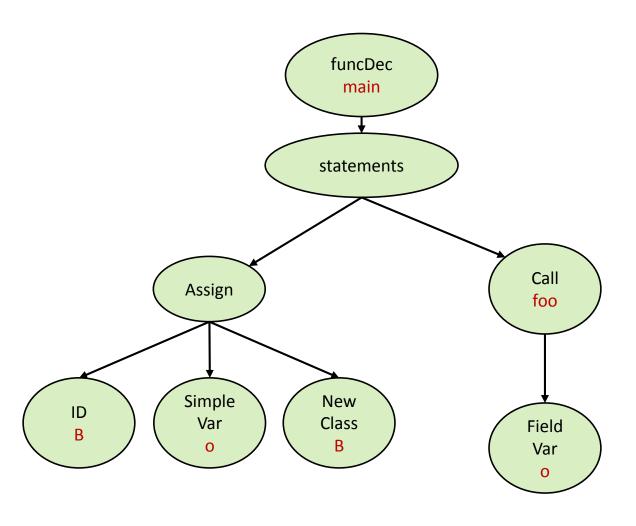
Invalid

```
class A { }
class B extends A { }
void foo(A a) { }
void main() {
  B o = new B;
  foo(o);
}
```

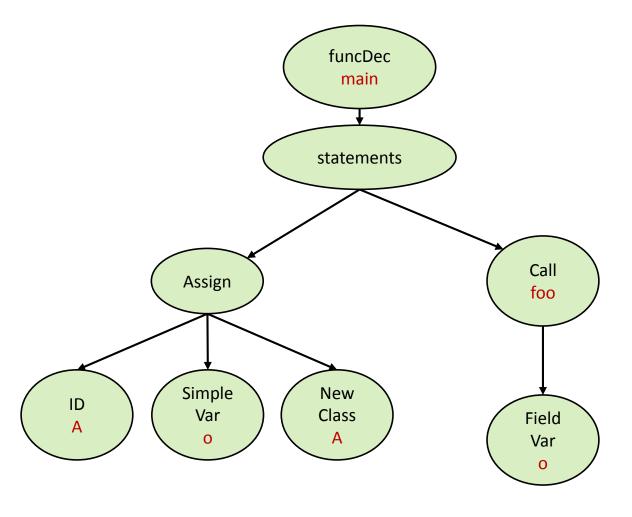


```
class A { }
class B extends A { }
void foo(A a) { }
void main() {
  B o = new B;
  foo(o);
}
```



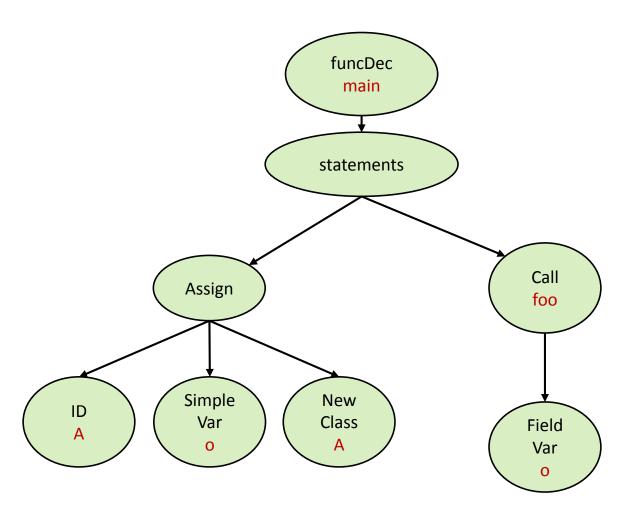


```
class A { }
class B extends A { }
void foo(B b) { }
void main() {
  A o = new A;
  foo(o);
}
```

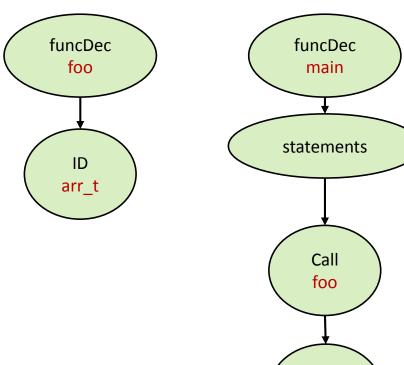


```
class A { }
class B extends A { }
void foo(B b) { }
void main() {
  A o = new A;
  foo(o);
}
```

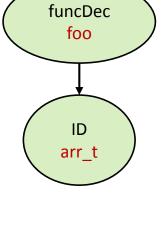
Invalid

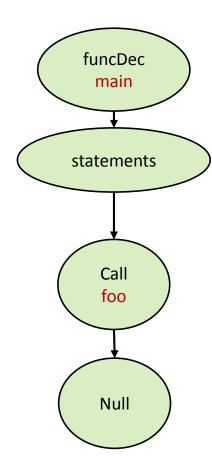


```
typedef int arr_t[];
void foo(arr_t a) { }
void main() {
  foo(null);
}
```



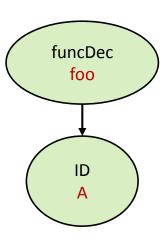
```
typedef int arr_t[];
void foo(arr_t a) { }
void main() {
  foo(null);
}
```

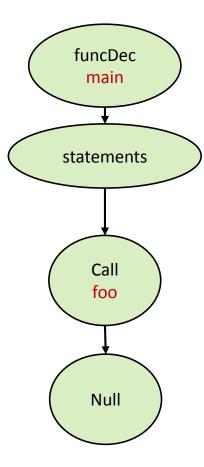






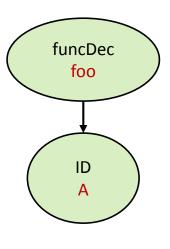
```
class A { };
void foo(A a) { }
void main() {
  foo(null);
}
```

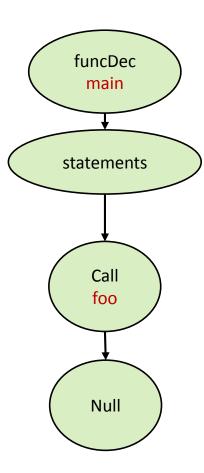




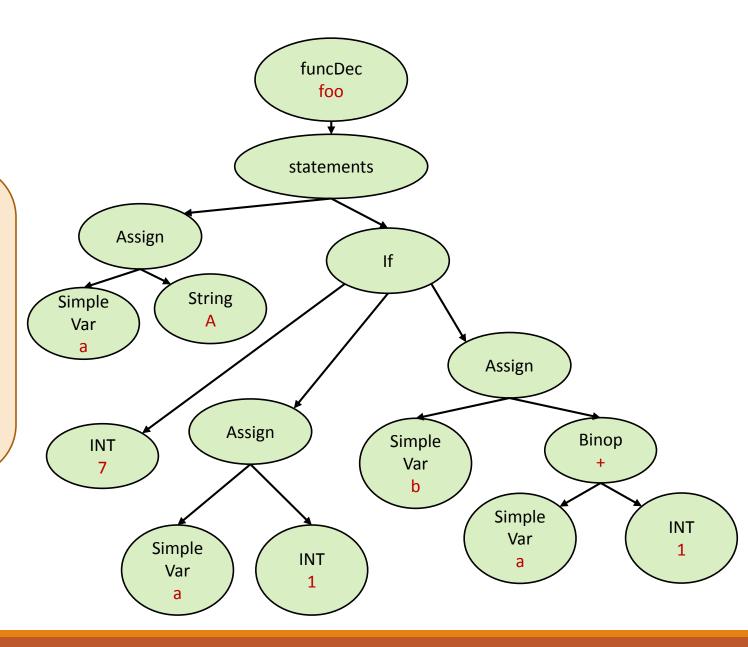
```
class A { };
void foo(A a) { }
void main() {
  foo(null);
}
```





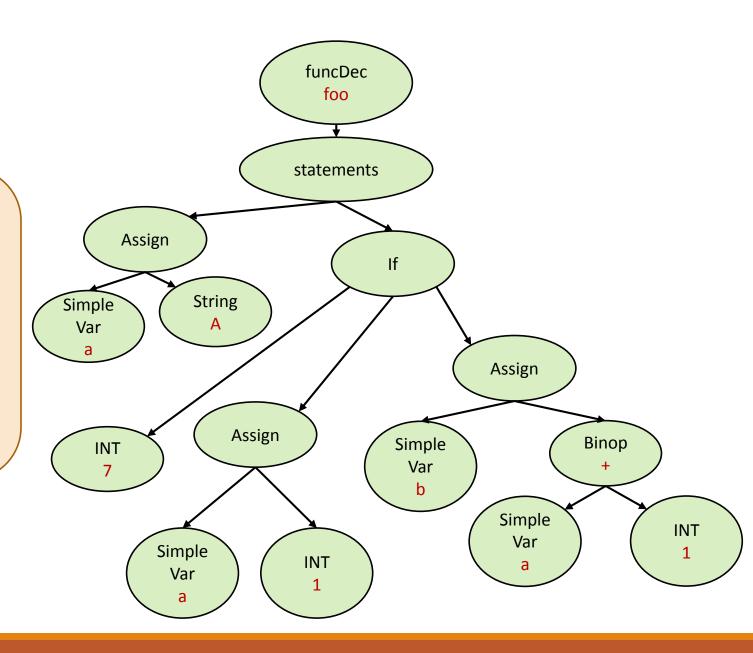


```
void foo(void) {
  string a = "A";
  if (7) {
    int a = 1;
    int b = a + 1;
  }
}
```

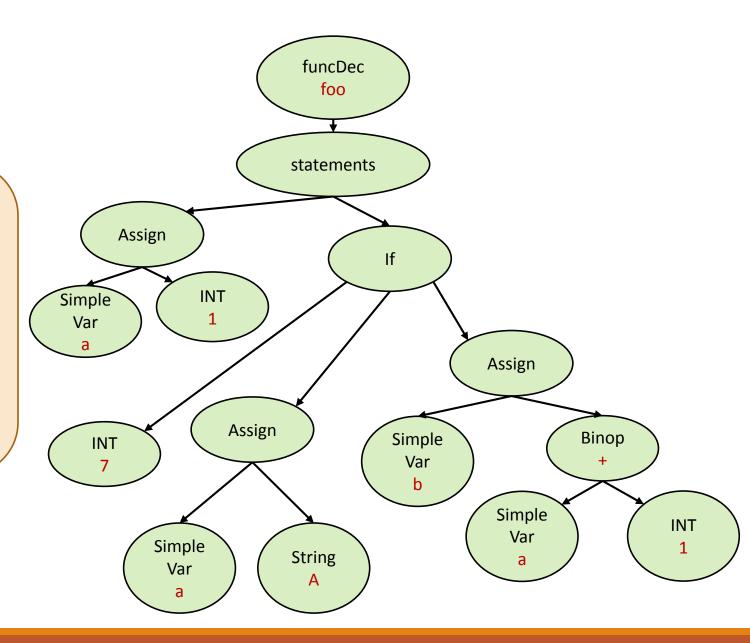


```
void foo(void) {
   string a = "A";
   if (7) {
      int a = 1;
      int b = a + 1;
   }
}
```

Valid

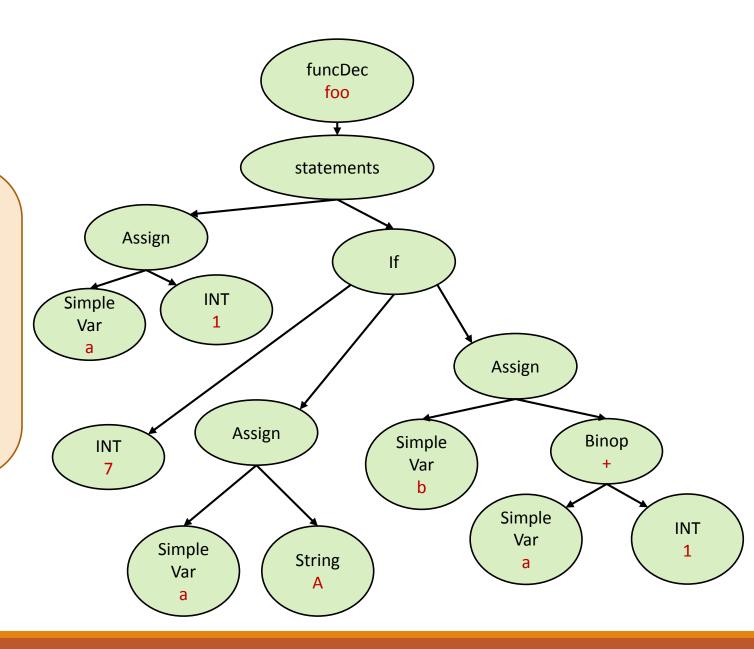


```
void foo(void) {
  int a = 1;
  if (7) {
    string a = "A";
    int b = a + 1;
  }
}
```

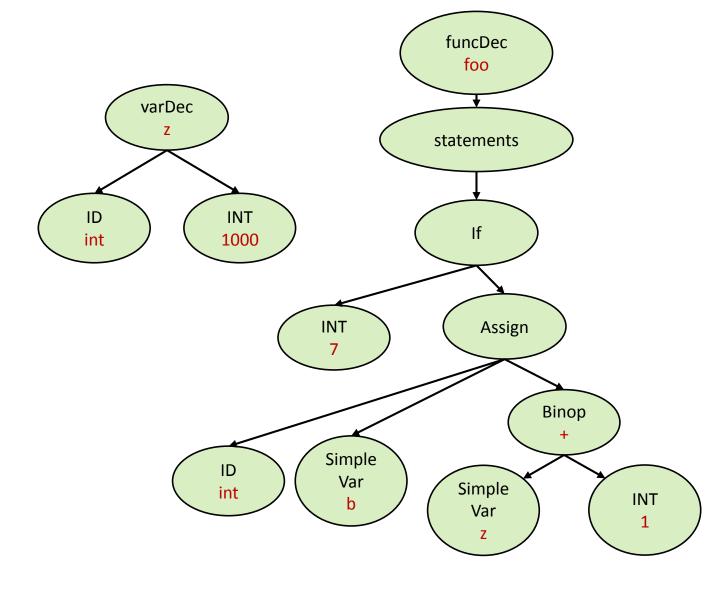


```
void foo(void) {
  int a = 1;
  if (7) {
    string a = "A";
    int b = a + 1;
  }
}
```

Invalid

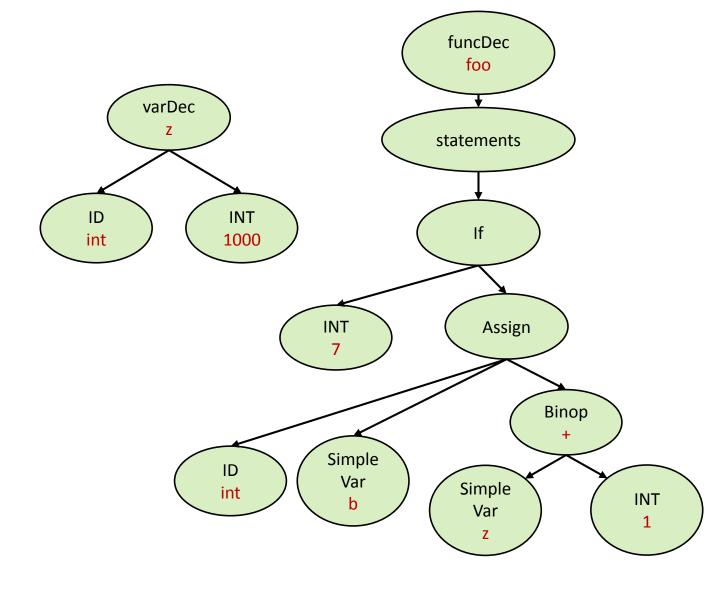


```
int z = 1000;
void foo(int z) {
   if (7) {
     int b = z + 1;
   }
}
```

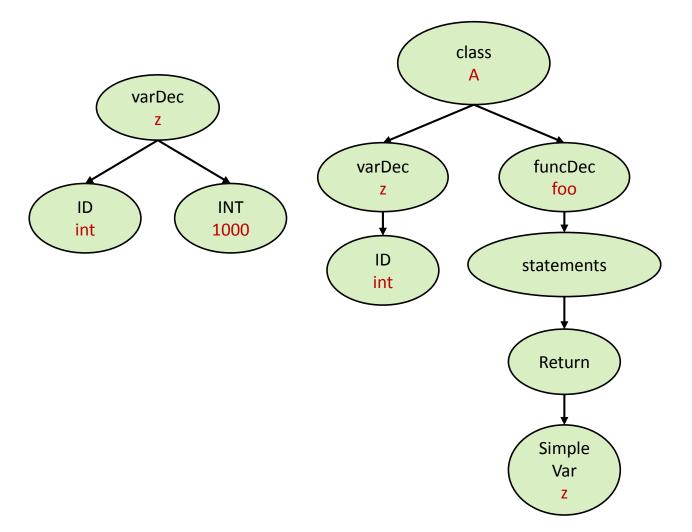


```
int z = 1000;
void foo(int z) {
   if (7) {
     int b = z + 1;
   }
}
```

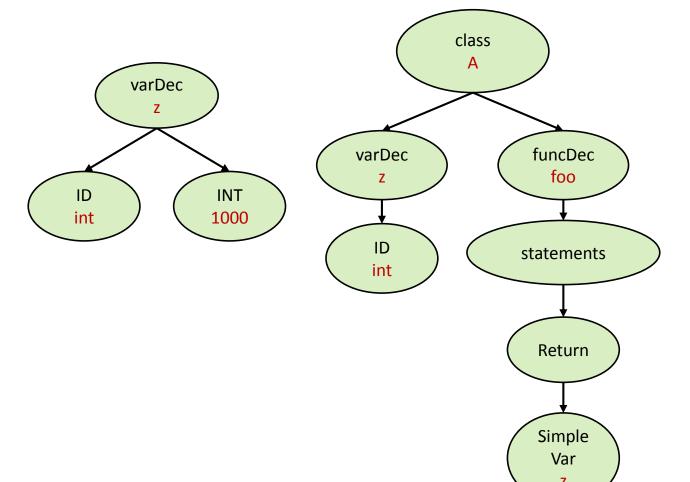
Valid



```
int z = 1000;
class A {
   int z;
   int foo() {
     return z;
   }
}
```

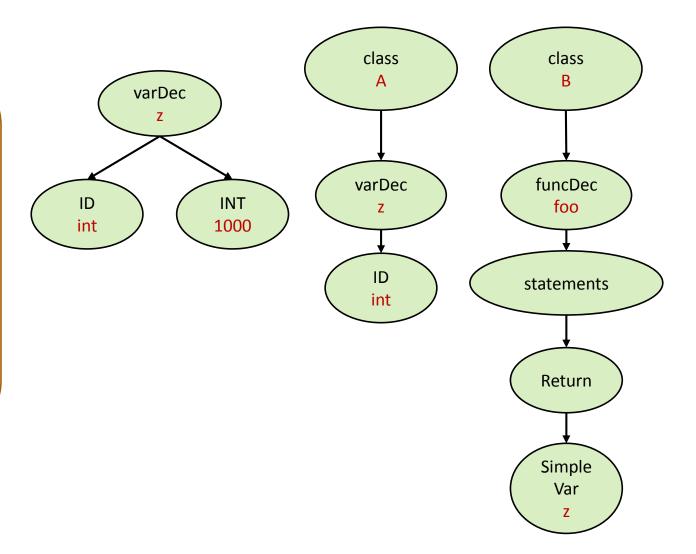


```
int z = 1000;
class A {
   int z;
   int foo() {
     return z;
   }
}
```

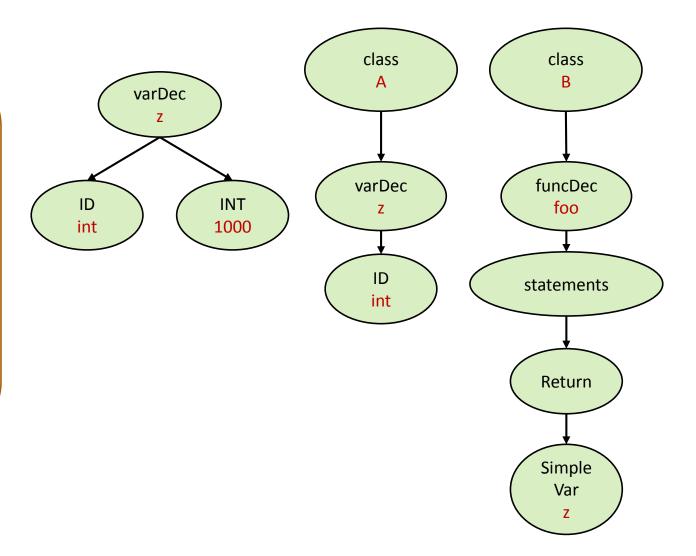


Valid

```
int z = 1000;
class A {
  int z;
class B extends A {
  int foo() {
    return z;
```



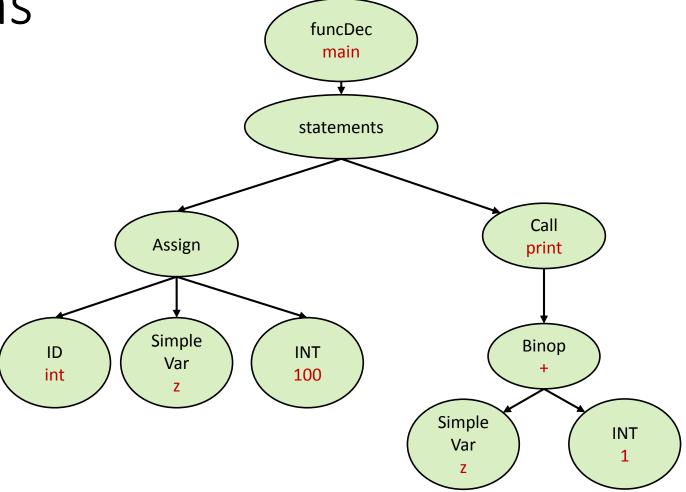
```
int z = 1000;
class A {
  int z;
class B extends A {
  int foo() {
    return z;
```



Valid

Library Functions

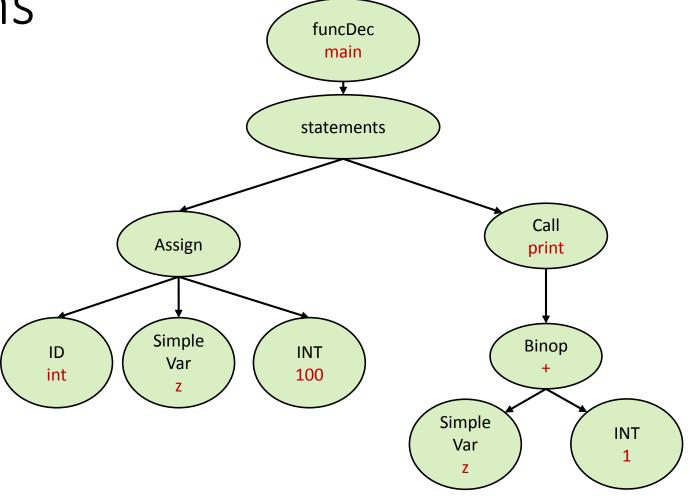
```
void main() {
  int z = 100;
  print(z + 1);
}
```



Library Functions

```
void main() {
  int z = 100;
  print(z + 1);
}
```

Valid



Implementation

The AST is traversed in a top-down manner

- Each AST node class, has a **visit** API:
 - Performs the relevant semantic checks
 - May call the visitors of the node's children
 - In the skeleton it's called *semantMe*
- The traversal starts from the root node

Implementation

```
Class ASTExpBinOp {
 public ASTExp left;
 public ASTExp right;
 public Type visit() {
    Type t1 = left.visit();
    Type t2 = right.visit();
    if (t1 != t2) {
     // error
```

Implementation

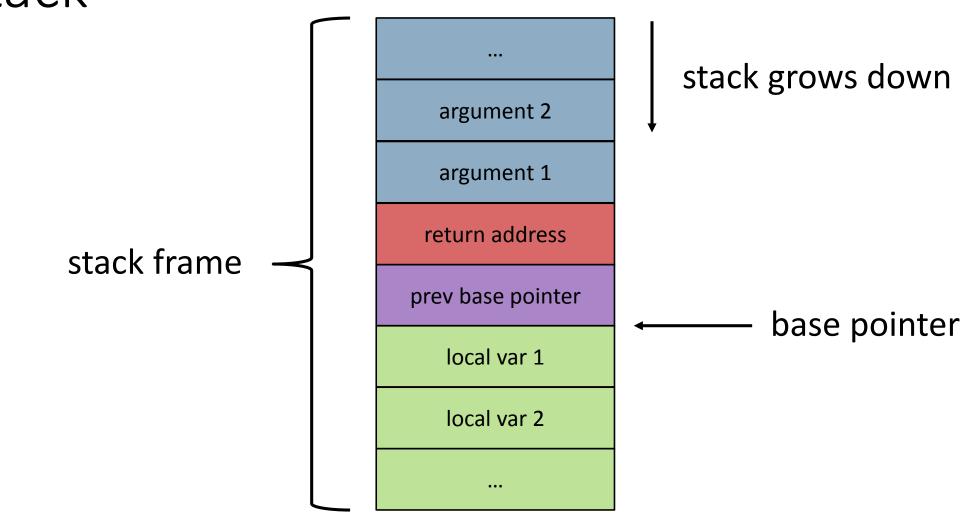
```
Class ASTStatmentList {
 public ASTStatement head;
 public ASTStatmentList tail;
 public Type visit() {
    if (head)
      head.visit();
    if (tail)
      tail.visit();
    return null;
```

AST Annotaations

AST Annotations

While analyzing the AST, we can extend it with useful information:

- Variable offsets
- Parameter offsets
- Class layout
- Type size



```
int f(int x, int y) {
   int z = x + y;
   return z;
}
int g() {
   int x = f(10, 20)
}
```

```
f:
push
       %ebp
       %esp,%ebp
mov
sub
       $0x10,%esp
       0x8(%ebp), %edx
mov
       0xc(%ebp), %eax
mov
      %edx, %eax
add
       %eax, -0x4(%ebp)
mov
       -0x4(%ebp), %eax
mov
leave
ret
g:
       $0x14
push
push
       $0x0a
       24 < g + 0xe >
call
```

argument 2

```
f:
push
      %ebp
       %esp,%ebp
mov
       $0x10,%esp
sub
       0x8(%ebp),%edx
mov
       0xc(%ebp), %eax
mov
      %edx,%eax
add
      %eax, -0x4 (%ebp)
mov
       -0x4(%ebp),%eax
mov
leave
ret
g:
```

••

push \$0x14
push \$0x0a
call 24 <g+0xe>

•••

argument 2

argument 1

```
f:
       %ebp
push
       %esp,%ebp
mov
       $0x10,%esp
sub
       0x8(%ebp),%edx
mov
       0xc(%ebp), %eax
mov
       %edx, %eax
add
       %eax, -0x4 (%ebp)
mov
       -0x4(%ebp),%eax
mov
leave
ret
g:
```

••

push \$0x14
push \$0x0a
call 24 <g+0xe>

•••

argument 2

argument 1

return address

```
f:
       %ebp
push
       %esp,%ebp
mov
       $0x10,%esp
sub
       0x8(%ebp),%edx
mov
       0xc(%ebp), %eax
mov
       %edx, %eax
add
       %eax, -0x4 (%ebp)
mov
       -0x4(%ebp),%eax
mov
leave
ret
```

g:

••

 push
 \$0x14

 push
 \$0x0a

call 24 < g+0xe>

...

argument 2

argument 1

return address

prev base pointer

f:

%ebp push %esp,%ebp mov \$0x10,%esp sub 0x8(%ebp),%edx mov 0xc(%ebp), %eax mov %edx, %eax add %eax, -0x4(%ebp) mov -0x4(%ebp),%eax mov leave ret

g:

••

push \$0x14
push \$0x0a
call 24 <g+0xe>

•••

argument 2

argument 1

return address

prev base pointer

base pointer ----

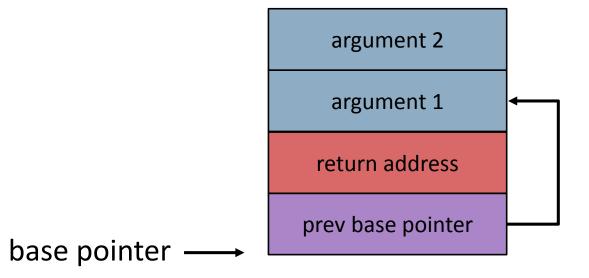
```
f:
push
       %ebp
       %esp,%ebp
mov
       $0x10,%esp
sub
       0x8(%ebp),%edx
mov
       0xc(%ebp), %eax
mov
       %edx, %eax
add
       %eax, -0x4(%ebp)
mov
       -0x4(%ebp), %eax
mov
leave
ret
```

g:

••

push \$0x14
push \$0x0a
call 24 <g+0xe>

•••



```
f:
push
       %ebp
       %esp,%ebp
mov
       $0x10,%esp
sub
       0x8(%ebp),%edx
mov
       0xc(%ebp), %eax
mov
       %edx, %eax
add
       %eax, -0x4(%ebp)
mov
       -0x4(%ebp), %eax
mov
leave
ret
g:
       $0x14
push
push
       $0x0a
       24 < g + 0xe >
call
```

argument 2 argument 1 return address prev base pointer base pointer ----

```
f:
push
       %ebp
       %esp,%ebp
mov
       $0x10,%esp
sub
       0x8(%ebp),%edx
mov
       0xc(%ebp),%eax
mov
       %edx,%eax
add
       %eax, -0x4(%ebp)
mov
       -0x4(%ebp), %eax
mov
leave
ret
g:
```

\$0x14 push push \$0x0a 24 < g + 0xe >call

argument 2

argument 1

return address

prev base pointer

local var 1

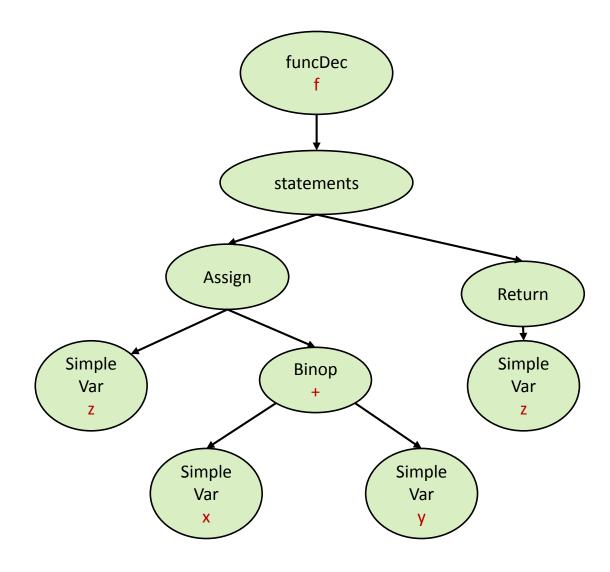
```
f:
push
       %ebp
       %esp,%ebp
mov
       $0x10,%esp
sub
       0x8(%ebp),%edx
mov
       0xc(%ebp), %eax
mov
       %edx, %eax
add
       %eax,-0x4(%ebp)
mov
       -0x4(%ebp), %eax
mov
leave
ret
g:
       $0x14
push
push
       $0x0a
       24 < g + 0xe >
call
```

Machine code does not contain names of:

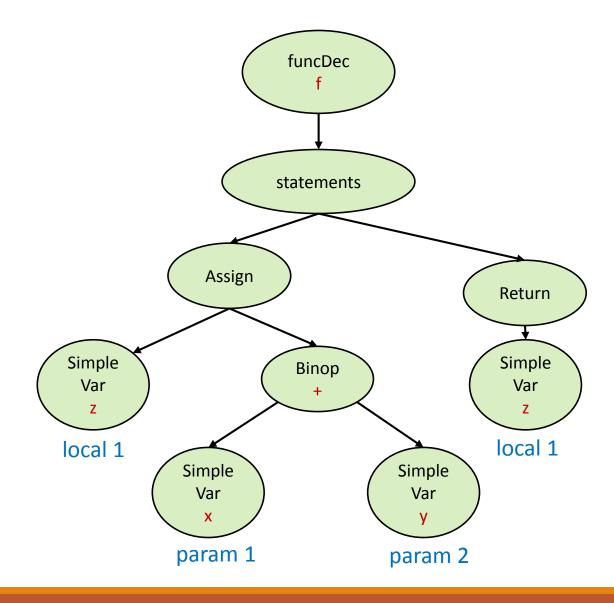
- Local variables
- Parameters

Instead, we use offsets relatively from the stack base pointer

```
int f(int x, int y) {
  int z = x + y;
  return z;
}
```



```
int f(int x, int y) {
   int z = x + y;
   return z;
}
```



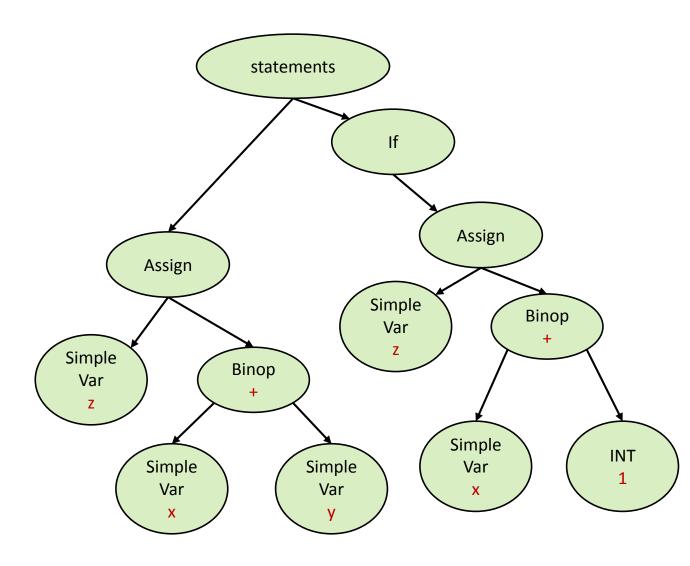
```
int f(int x, int y) {
  int z = x + y;
  return z;
}

x ⇒ param 1 ⇒ 0x08

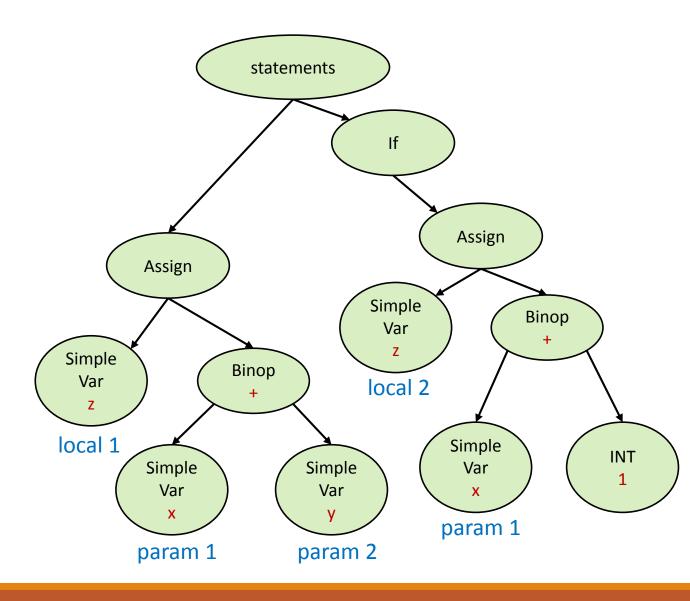
y ⇒ param 2 ⇒ 0x0c

z ⇔ local 1 ⇒ -0x04
```

```
void f(int x, int y) {
  int z = x + y;
  if (z > 1) {
    int z = x + 1;
  }
}
```

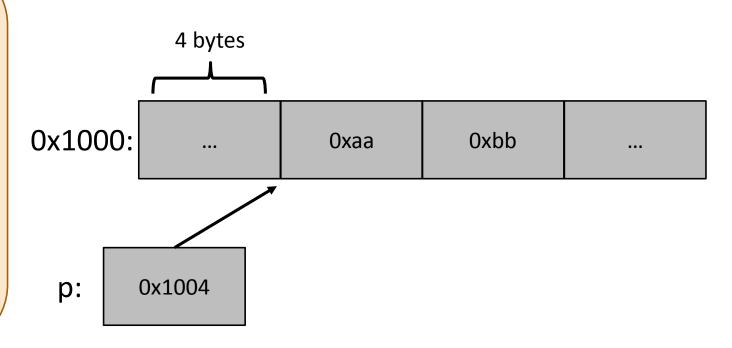


```
void f(int x, int y) {
  int z = x + y;
  if (z > 1) {
    int z = x + 1;
  }
}
```



```
class Point {
  int x;
  int y;
void f(Point p) {
 p.x = 0xaa;
  p.y = 0xbb;
```

```
class Point {
  int x;
  int y;
void f(Point p) {
  p.x = 0xaa;
  p.y = 0xbb;
```



```
class Point {
  int x;
  int y;
void f(Point p) {
  p.x = 0xaa;
  p.y = 0xbb;
```

```
f:
       %ebp
push
       %esp, %ebp
MOV
       0x8(%ebp), %eax
MOV
       $0xaa, (%eax)
movl
       0x8(%ebp), %eax
mov
       $0xbb,0x4(%eax)
movl
       %ebp
pop
ret
```

```
class Point {
  int x;
  int y;
void f(Point p) {
  p.x = 0xaa;
  p.y = 0xbb;
```

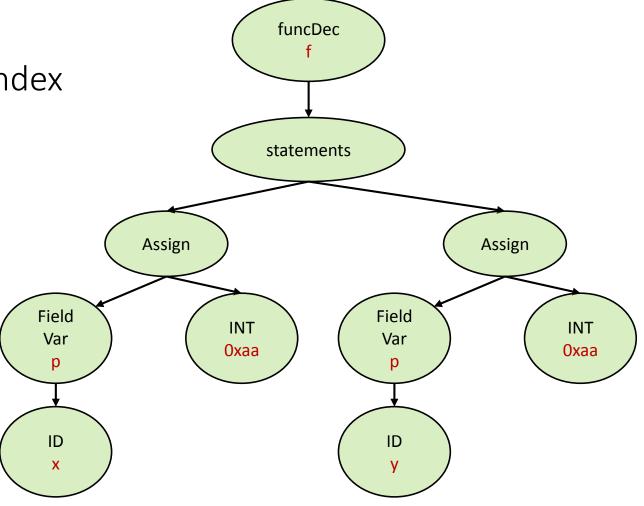
```
f:
       %ebp
push
       %esp, %ebp
MOV
       0x8(%ebp), %eax
mov
movl
       $0xaa, (%eax)
       0x8(%ebp), %eax
mov
       $0xbb, 0x4(%eax)
movl
       %ebp
pop
ret
```

```
class Point {
  int x;
  int y;
void f(Point p) {
  p.x = 0xaa;
  p.y = 0xbb;
```

```
f:
       %ebp
push
       %esp, %ebp
MOV
       0x8(%ebp), %eax
MOV
       $0xaa, (%eax)
movl
       0x8(%ebp), %eax
mov
       $0xbb, 0x4 (%eax)
movl
       %ebp
pop
ret
```

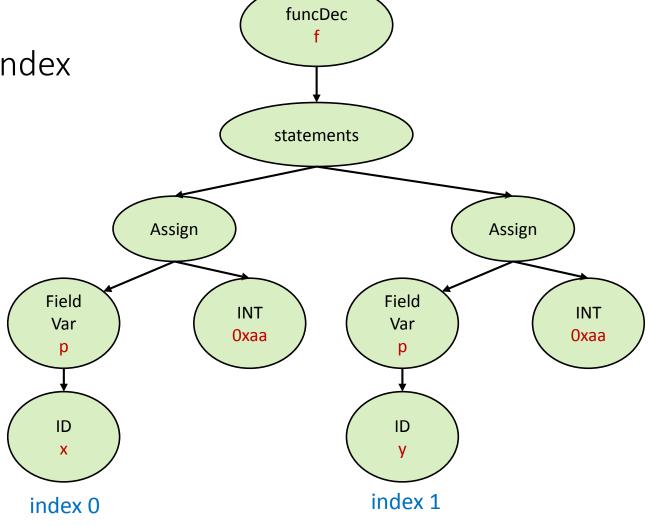
Each class field should have an index

```
class Point {
  int x;
  int y;
}
void f(Point p) {
  p.x = 0xaa;
  p.y = 0xbb;
}
```



Each class field should have an index

```
class Point {
  int x;
  int y;
}
void f(Point p) {
  p.x = 0xaa;
  p.y = 0xbb;
}
```



```
class Point {
  int x;
  int y;
}
void f(Point p) {
  Point p = new Point;
}
```

```
f:
push
       %ebp
       %esp, %ebp
mov
sub
       $0x18,%esp
       $0xc, %esp
sub
push
       $0x8
call 56 < foo + 0xc >
       $0x10,%esp
add
       %eax, -0xc(%ebp)
MOV
leave
ret
```

```
class Point {
  int x;
  int y;
}
void f(Point p) {
  Point p = new Point;
}
```

```
f:
push
       %ebp
       %esp, %ebp
mov
sub
       $0x18,%esp
       $0xc, %esp
sub
push
       $0x8
call 56 < foo + 0xc >
       $0x10,%esp
add
       %eax, -0xc(%ebp)
mov
leave
ret
```

```
class Point {
  int x;
  int y;
  string name;
}
void f(Point p) {
  Point p = new Point;
}
```

```
f:
push
       %ebp
       %esp,%ebp
mov
       $0x18,%esp
sub
       $0xc, %esp
sub
push
call 56 < foo + 0xc >
       $0x10,%esp
add
       %eax, -0xc(%ebp)
MOV
leave
ret
```

```
class Point {
  int x;
  int y;
  string name;
}
void f(Point p) {
  Point p = new Point;
}
```

```
f:
push
       %ebp
       %esp, %ebp
mov
       $0x18,%esp
sub
       $0xc, %esp
sub
push
       $0xc
call 56 < foo + 0xc >
       $0x10,%esp
add
       %eax, -0xc(%ebp)
mov
leave
ret
```