

```
print(a + { a += 1; a })
```









```
// a = 2
```

```
print(1 + 2)
```

```
// -> 4
```





```
a = 1
```

```
def f():  
    global a  
    a += 1  
    return a
```

```
print(a + f())
```





```
var a = 1
```

```
print(a + {  
    a += 1;  
    a  
})
```

```
a = 1
```

```
def f():  
    global a  
    a += 1  
    return a
```

```
print(a + f())
```



```
a = 1

def f():
    global a
    a += 1
    return a

print(a + f())
```

```
Windows PowerShell x + v
PS C:\dev\kibi\bug> python.exe .\bug.py
3
```



```
local a = 1
```

```
function f()  
    a = a + 1  
    return a  
end
```

```
print(a + f())
```



```
local a = 1

function f()
    a = a + 1
    return a
end

print(a + f())
```

```
Windows PowerShell x + v
PS C:\dev\kibi\bug> lua54.exe .\bug.lua
4
```



```
a = 1

def f():
    global a
    a += 1
    return a

print(a + f())
```

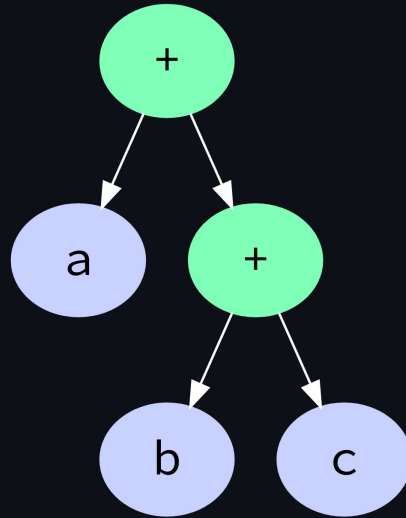
```
Windows PowerShell x + v
PS C:\dev\kibi\bug> python.exe .\bug.py
3
```

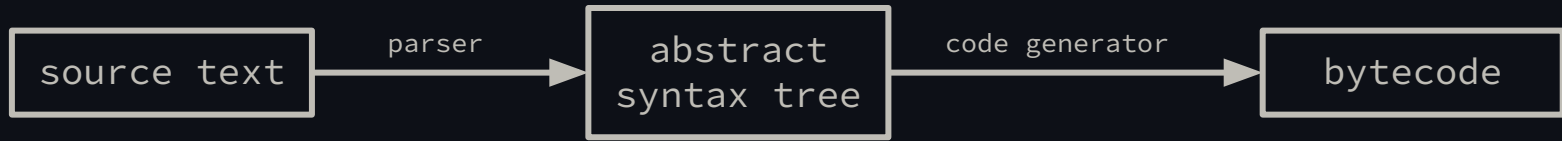
source text

"a + (b + c)"

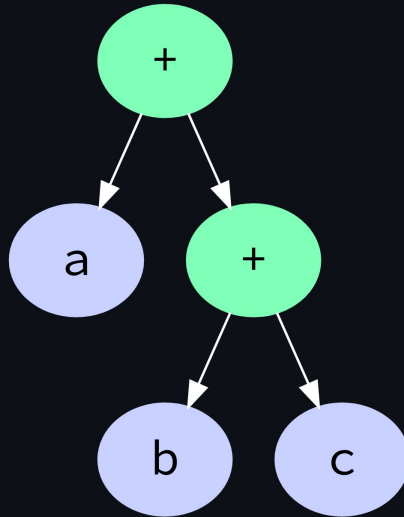


"a + (b + c)"



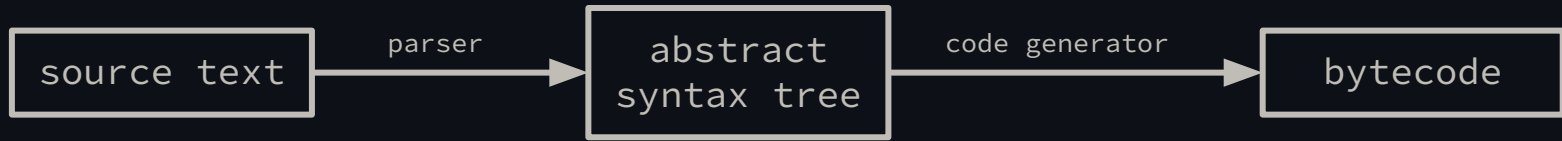


"a + (b + c)"

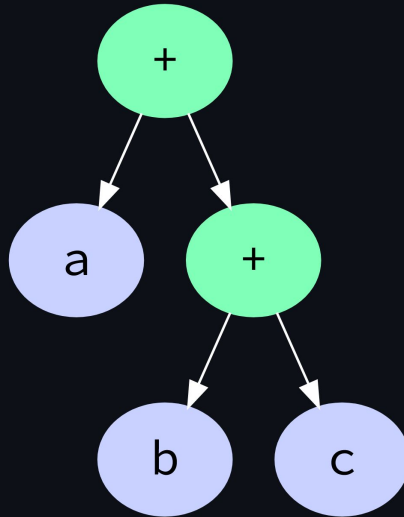


```
LOAD_FAST 0 (a)
LOAD_FAST 1 (b)
LOAD_FAST 2 (c)
BINARY_ADD
BINARY_ADD
```

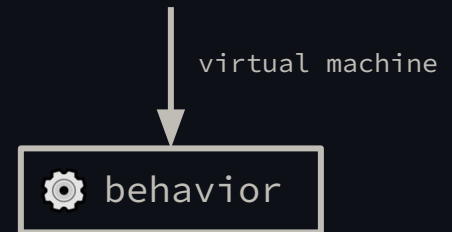




"a + (b + c)"



```
LOAD_FAST 0 (a)
LOAD_FAST 1 (b)
LOAD_FAST 2 (c)
BINARY_ADD
BINARY_ADD
```



BINARY\_ADD

20

22

foo

69

---

BINARY\_ADD

22

20

foo

69

---

BINARY\_ADD

42

foo

69

---

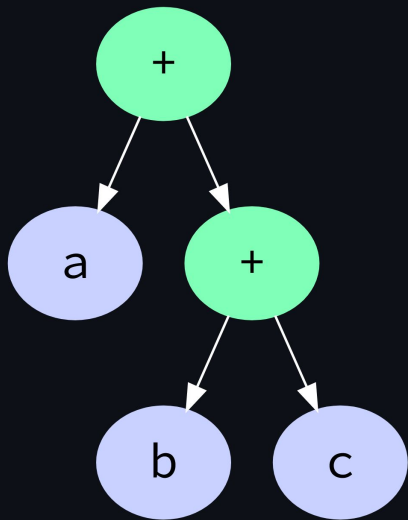
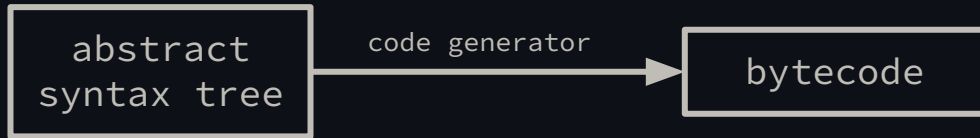
BINARY\_ADD

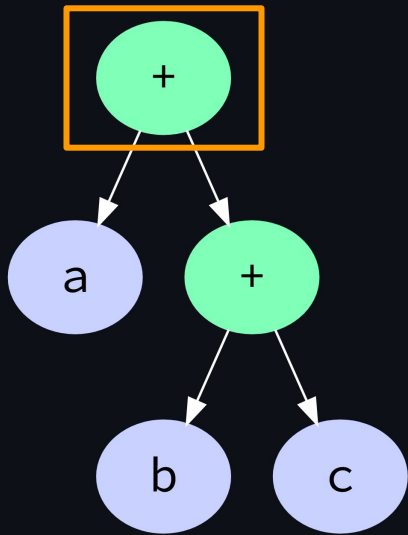
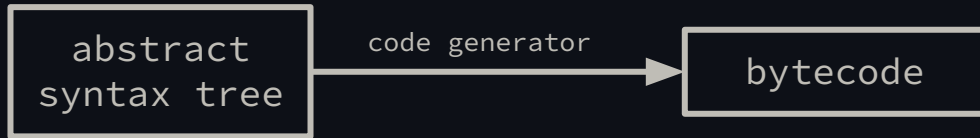
42

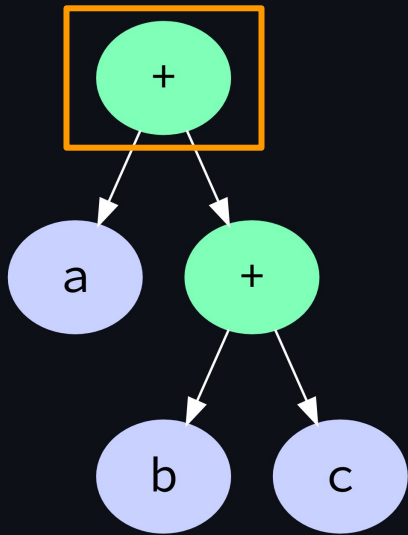
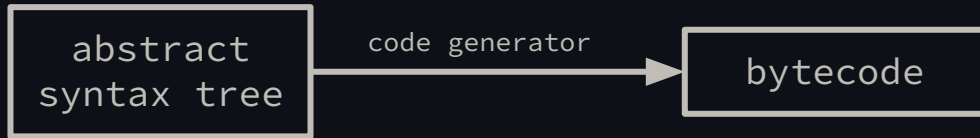
foo

69

---

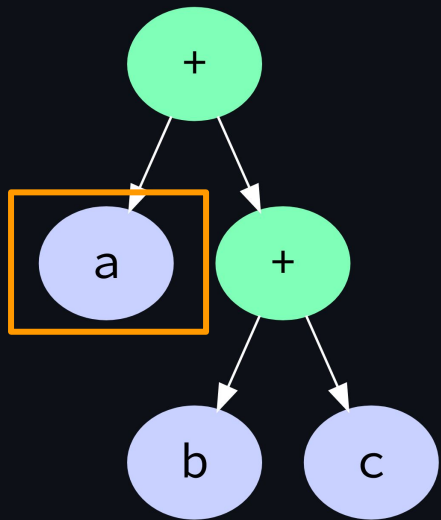
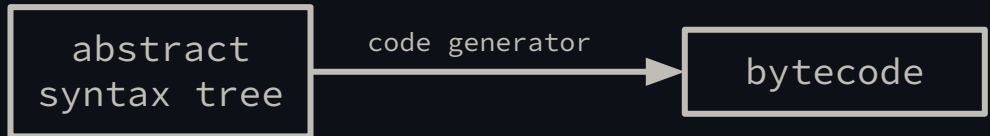




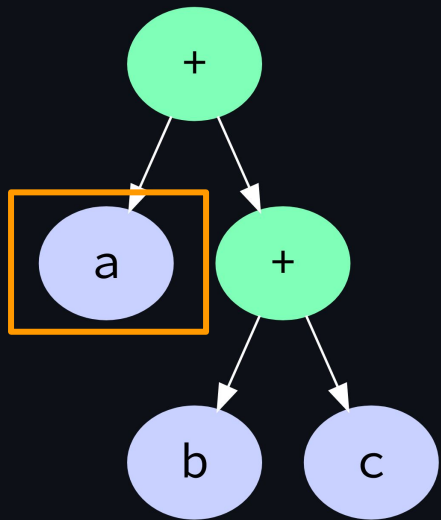
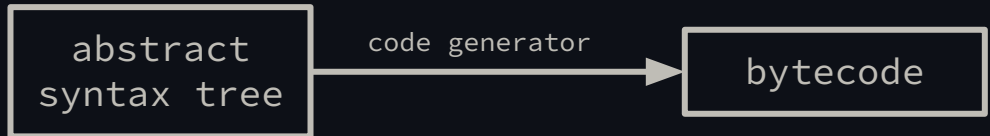


BINARY\_ADD



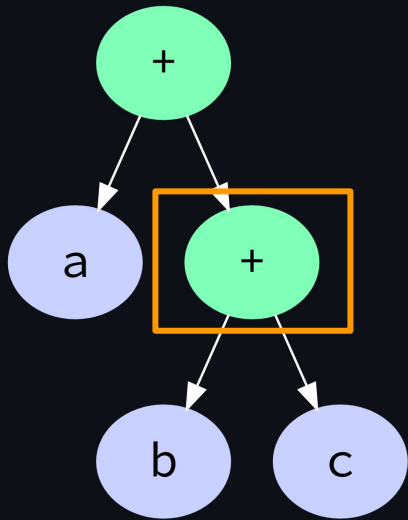
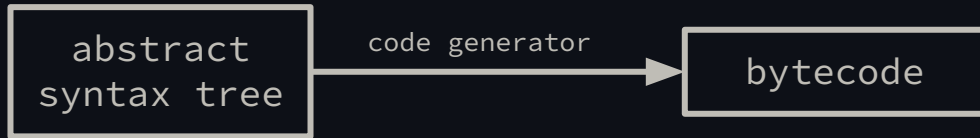


BINARY\_ADD



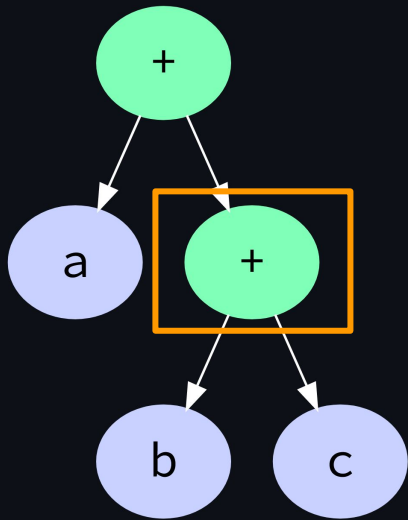
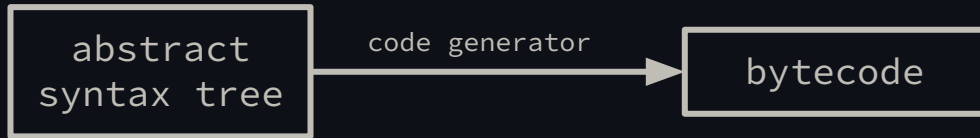
LOAD\_FAST 0 (a)

BINARY\_ADD



LOAD\_FAST 0 (a)

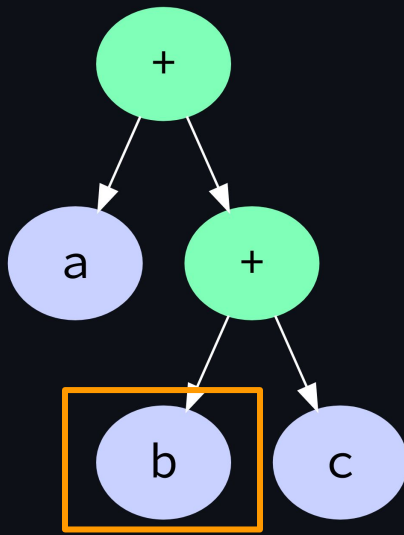
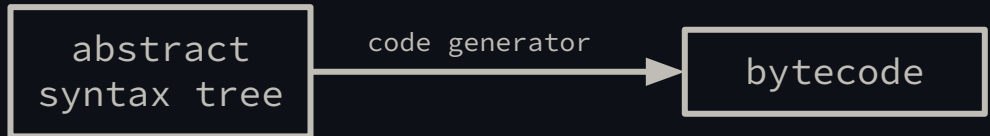
BINARY\_ADD



LOAD\_FAST 0 (a)

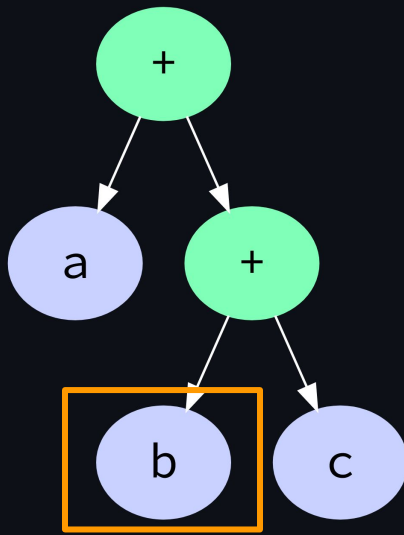
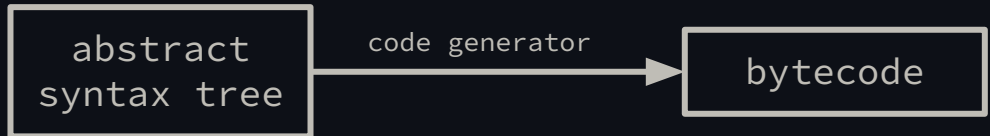
BINARY\_ADD

BINARY\_ADD



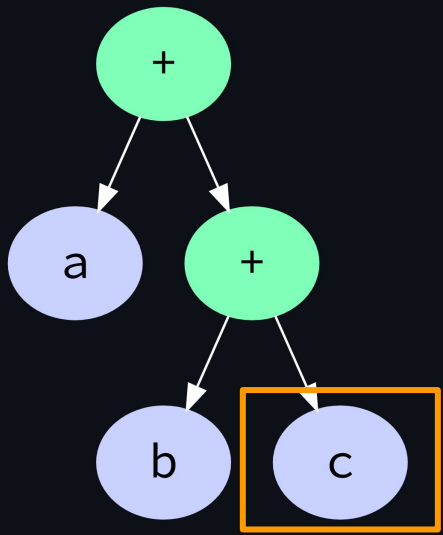
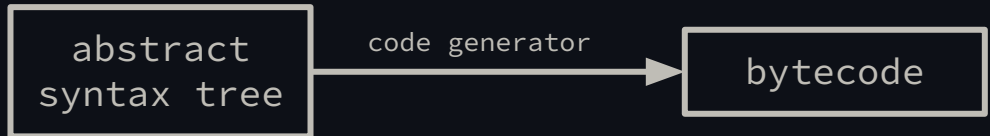
LOAD\_FAST 0 (a)

BINARY\_ADD  
BINARY\_ADD

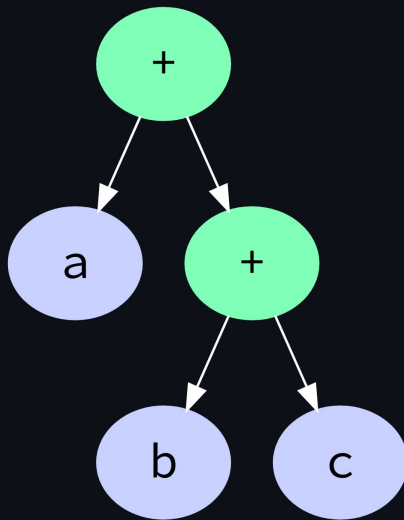
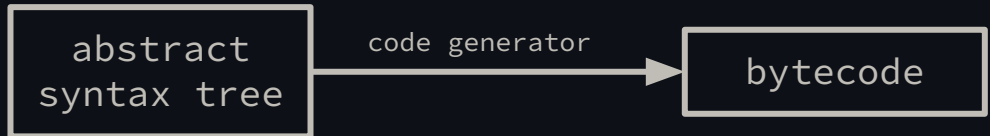


LOAD\_FAST 0 (a)  
LOAD\_FAST 1 (b)

BINARY\_ADD  
BINARY\_ADD



```
LOAD_FAST 0 (a)
LOAD_FAST 1 (b)
LOAD_FAST 2 (c)
BINARY_ADD
BINARY_ADD
```



```
LOAD_FAST 0 (a)
LOAD_FAST 1 (b)
LOAD_FAST 2 (c)
BINARY_ADD
BINARY_ADD
```



```
5640     static int
5641     compiler_visit_expr1(struct compiler *c, expr_ty e)
5642     {
5643         location loc = LOC(e);
5644         switch (e->kind) {
5652             case BinOp_kind:
5653                 VISIT(c, expr, e->v.BinOp.left);
5654                 VISIT(c, expr, e->v.BinOp.right);
5655                 ADDOP_BINARY(c, loc, e->v.BinOp.op);
5656                 break;
5727             case Constant_kind:
5728                 ADDOP_LOAD_CONST(c, loc, e->v.Constant.value);
5729                 break;
```

```
a = 1
```

```
def f():  
    global a  
    a += 1  
    return a
```

```
print(a + f())
```

```
a = 1
```

```
def f():  
    global a  
    a += 1  
    return a
```

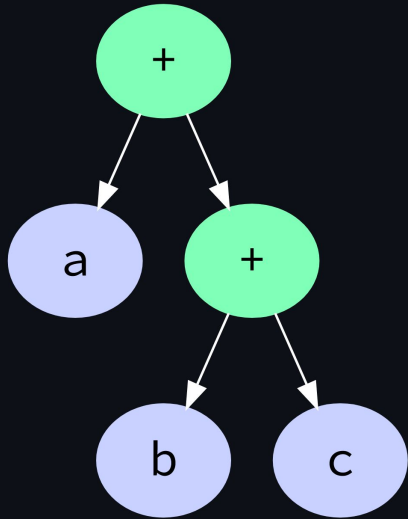
```
print(a + f())
```

```
a = 1
```

```
def f():  
    global a  
    a += 1  
    return a
```

```
print(a + f())
```

LOAD_GLOBAL	1	(a)
LOAD_GLOBAL	2	(f)
CALL_FUNCTION	0	
BINARY_ADD		



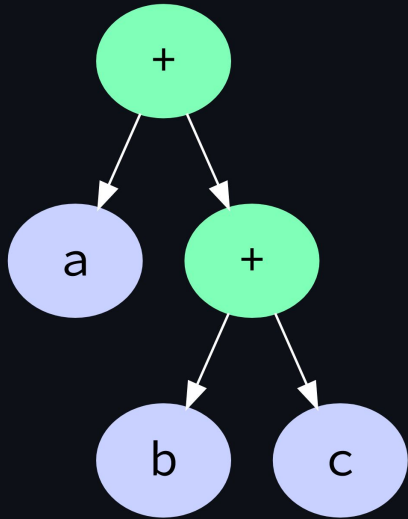
LOAD\_FAST 0 (a)

LOAD\_FAST 1 (b)

LOAD\_FAST 2 (c)

BINARY\_ADD

BINARY\_ADD



LOAD\_FAST 0 (a)

LOAD\_FAST 1 (b)

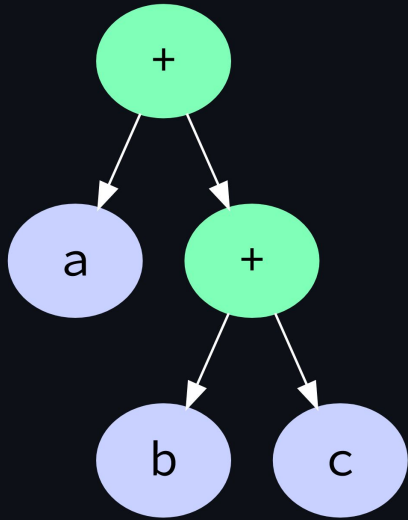
LOAD\_FAST 2 (c)

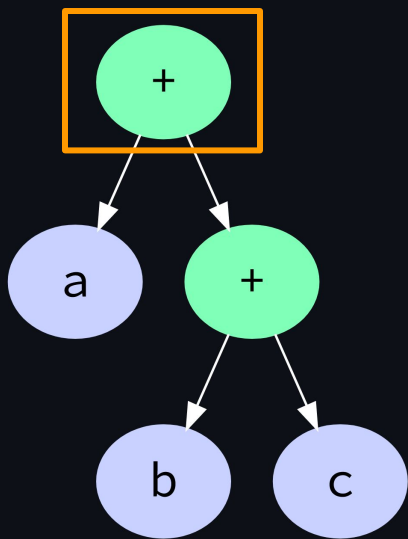
BINARY\_ADD

BINARY\_ADD

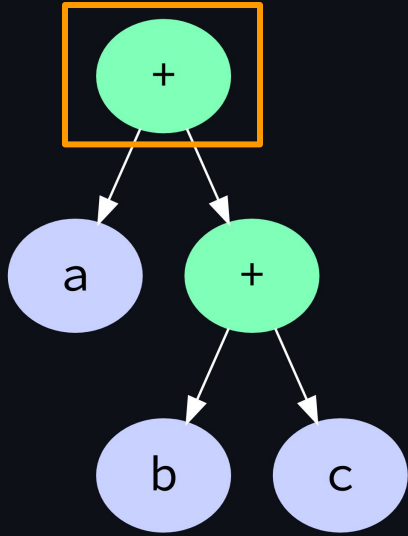
ADD r3, r1(b), r2(c)

ADD r4, r0(a), r3

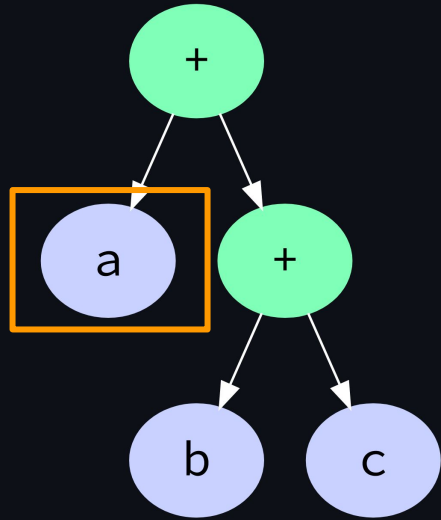




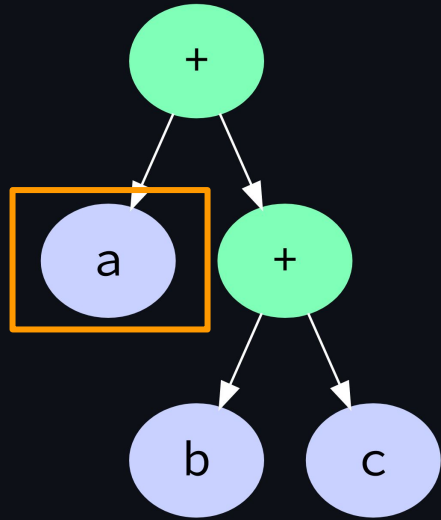




ADD r?, r?, r?

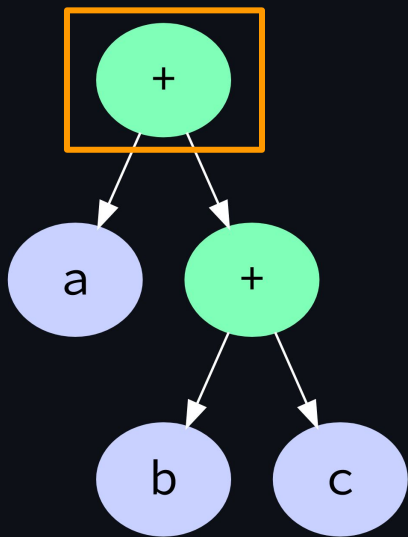


ADD r?, r?, r?



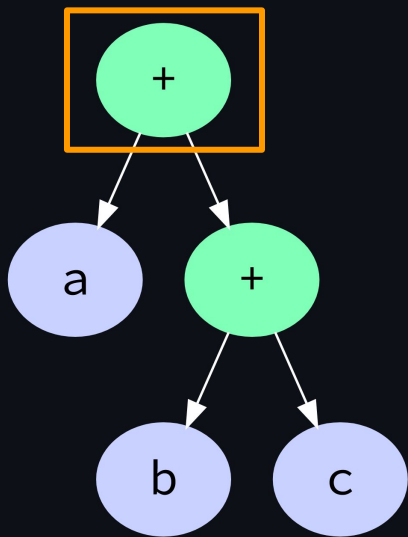
a: r0  
b: r1  
c: r2

ADD r?, r?, r?



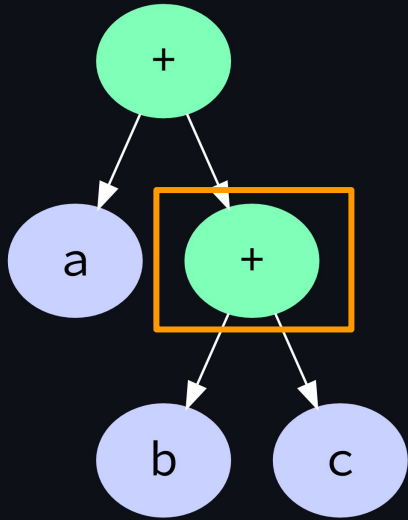
a: r0  
b: r1  
c: r2

ADD r?, r?, r?



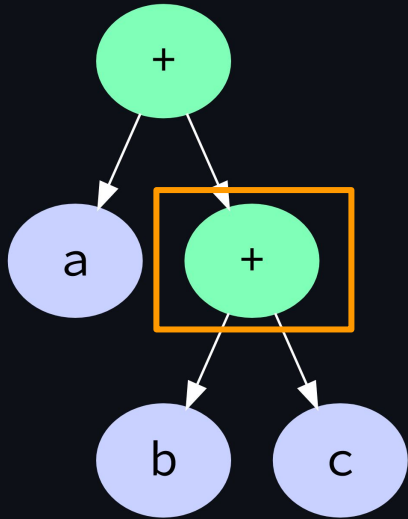
a: r0  
b: r1  
c: r2

ADD r?, r0(a), r?



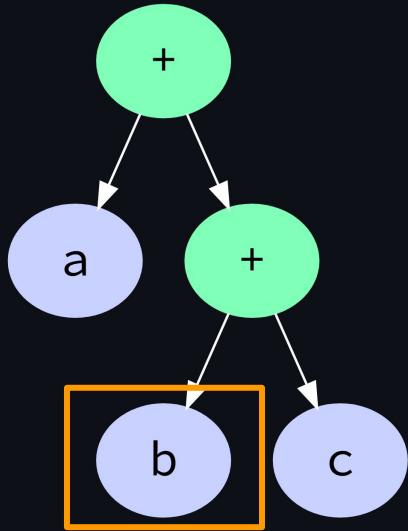
a: r0  
b: r1  
c: r2

ADD r?, r0(a), r?



a: r0  
b: r1  
c: r2

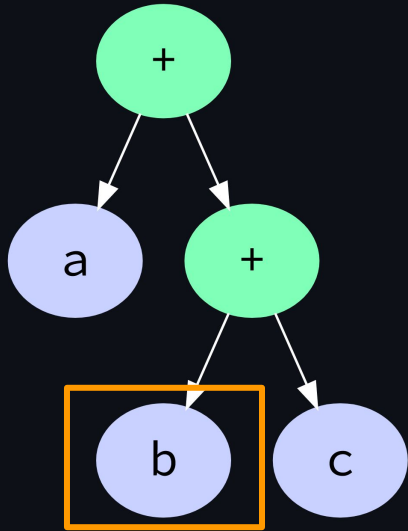
ADD r?, r?, r?  
ADD r?, r0(a), r?



a: r0  
b: r1  
c: r2

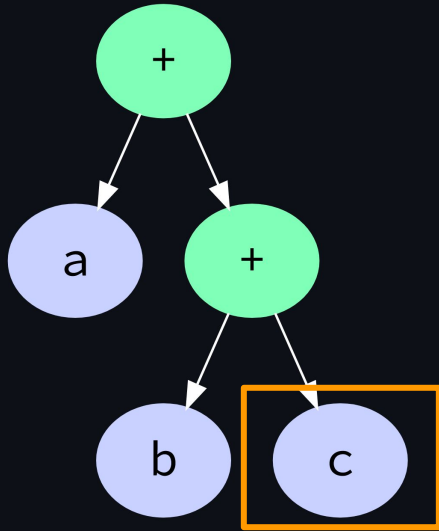
ADD r?, r?, r?  
ADD r?, r0(a), r?





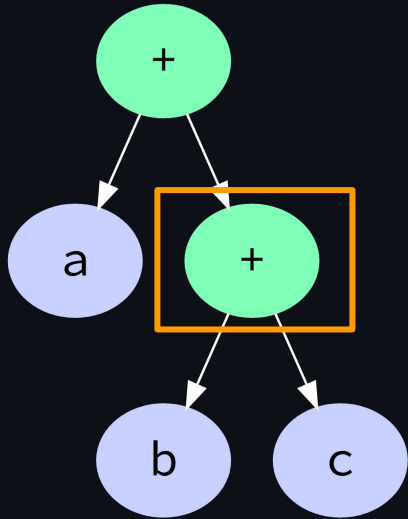
a: r0  
b: r1  
c: r2

ADD r?, r1(b), r?  
ADD r?, r0(a), r?



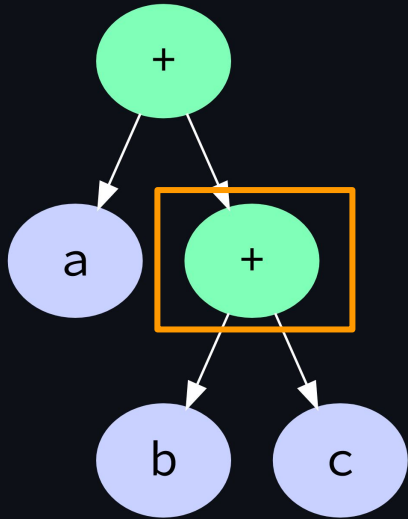
a: r0  
b: r1  
c: r2

ADD r?, r1(b), r2(c)  
ADD r?, r0(a), r?



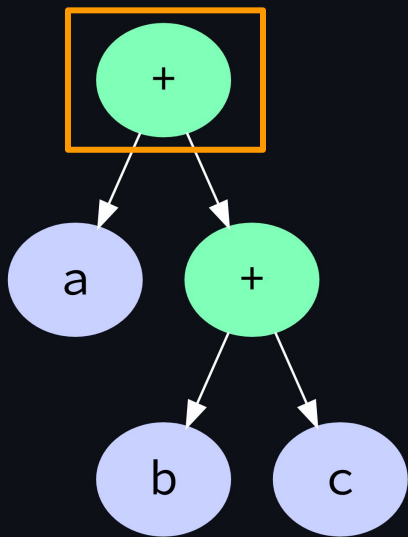
a: r0  
b: r1  
c: r2

ADD r?, r1(b), r2(c)  
ADD r?, r0(a), r?



a: r0  
b: r1  
c: r2

ADD r3, r1(b), r2(c)  
ADD r?, r0(a), r?



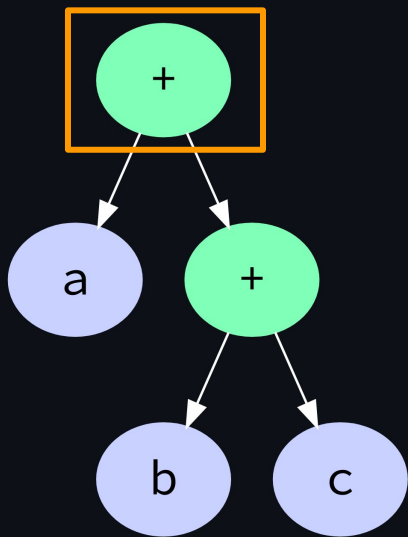
a: r0

b: r1

c: r2

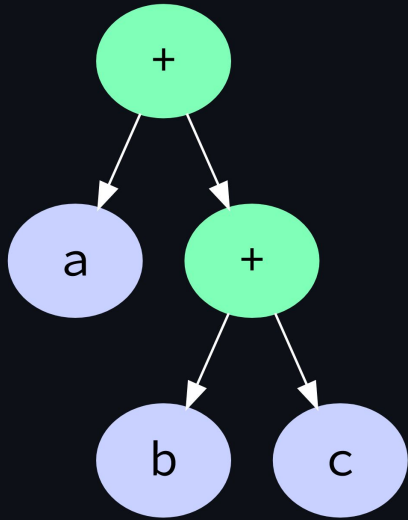
ADD r3, r1(b), r2(c)

ADD r?, r0(a), r3



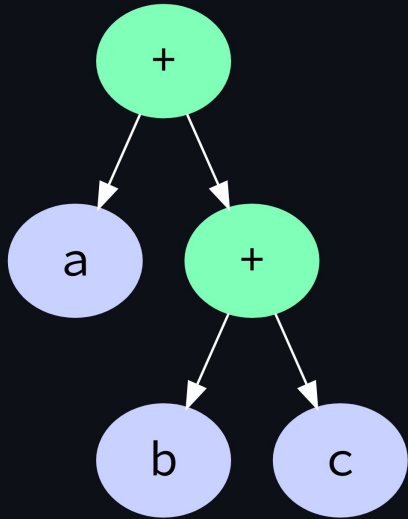
a: r0  
b: r1  
c: r2

```
ADD r3, r1(b), r2(c)  
ADD r4, r0(a), r3
```



a: r0  
b: r1  
c: r2

```
ADD r3, r1(b), r2(c)  
ADD r4, r0(a), r3
```



LOAD\_FAST 0 (a)

LOAD\_FAST 1 (b)

LOAD\_FAST 2 (c)

BINARY\_ADD

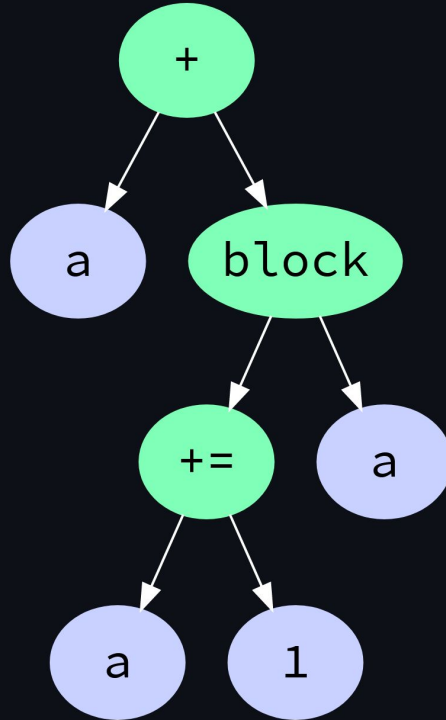
BINARY\_ADD

ADD r3, r1(b), r2(c)

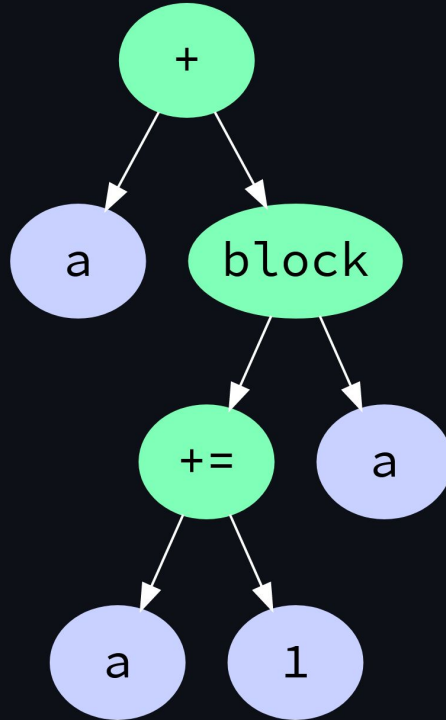
ADD r4, r0(a), r3



```
a + { a += 1; a }
```



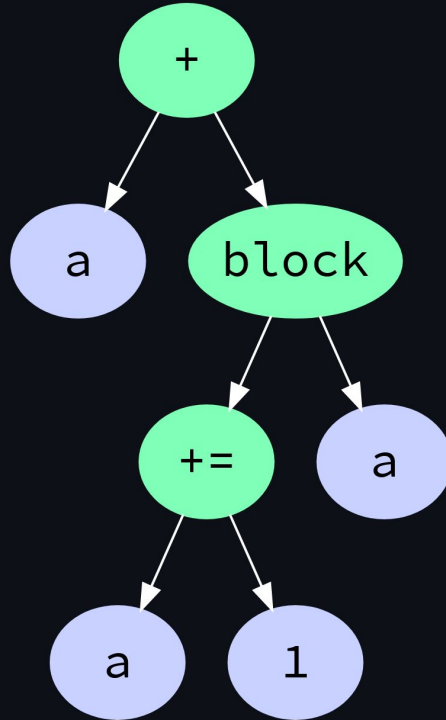
```
a + { a += 1; a }
```



```
ADD r0(a), r0(a), #1  
ADD r1, r0(a), r0(a)
```

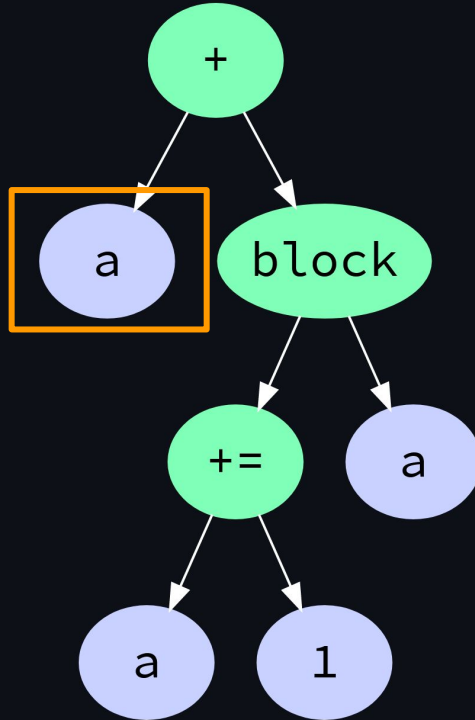


```
a + { a += 1; a }
```



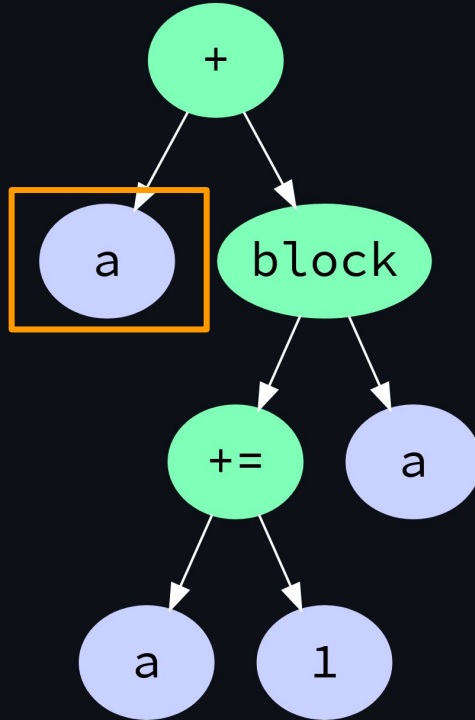
```
ADD r0(a), r0(a), #1  
ADD r1, r0(a), r0(a)
```

```
a + { a += 1; a }
```

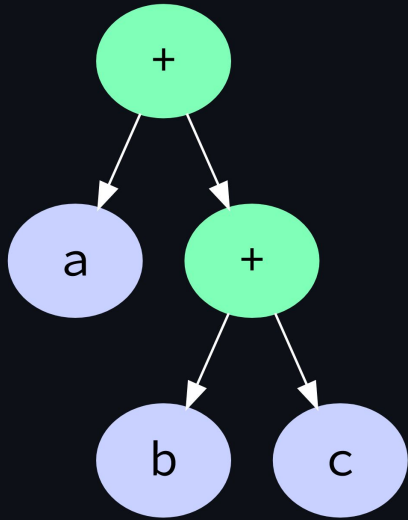


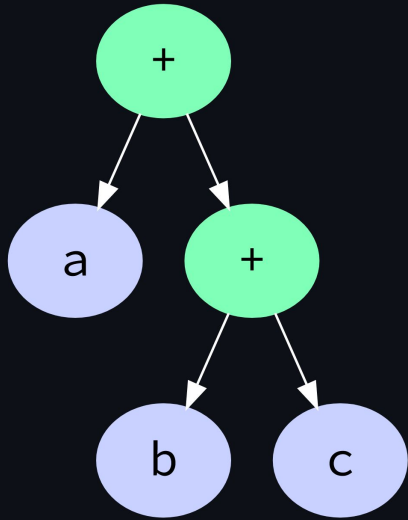
```
ADD r0(a), r0(a), #1  
ADD r1, r0(a), r0(a)
```

`a + { a += 1; a }`



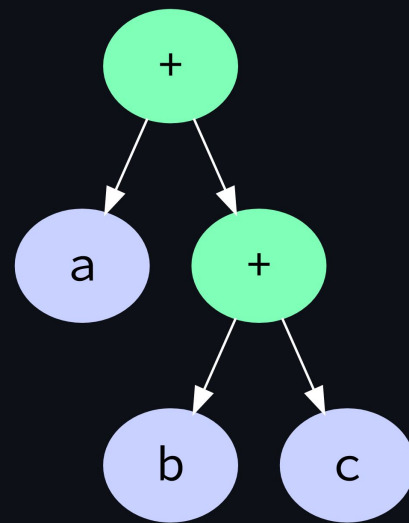
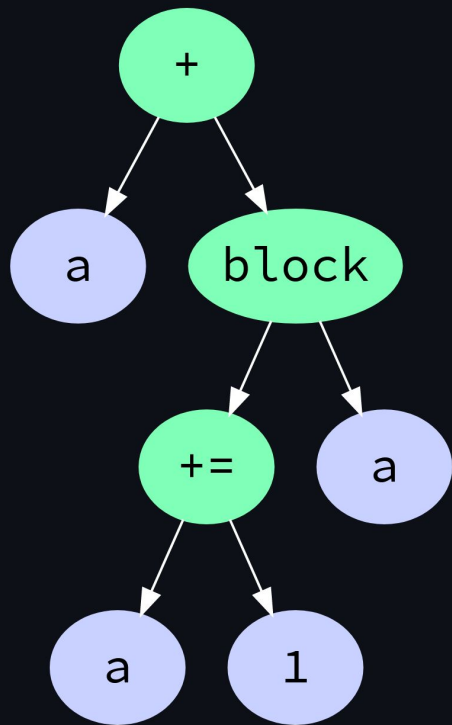
```
COPY r1, r0(a)
ADD r0(a), r0(a), #1
ADD r1, r1, r0(a)
```

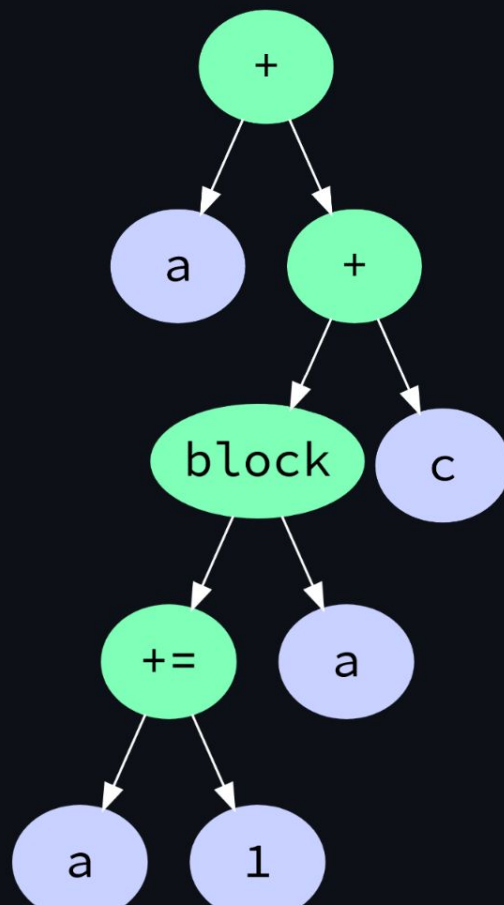
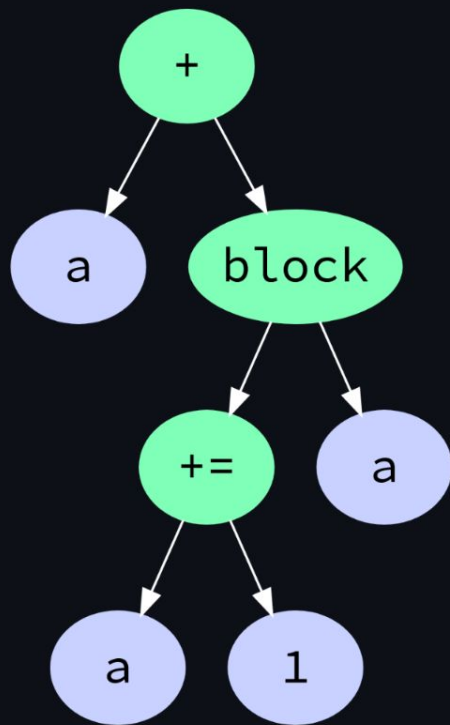




```
COPY r3, r0(a)
COPY r4, r1(b)
COPY r5, r2(c)
ADD r4, r4, r5
ADD r3, r3, r4
```





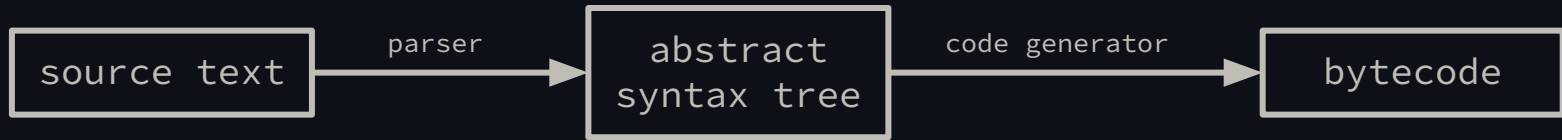




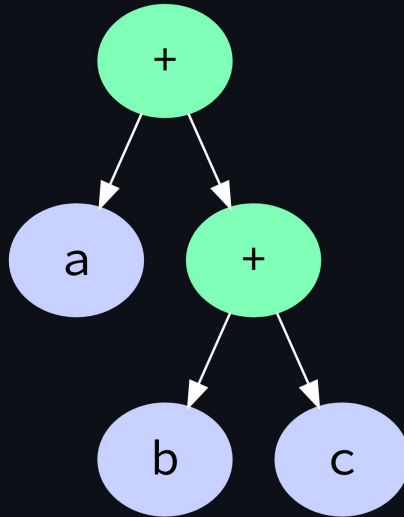




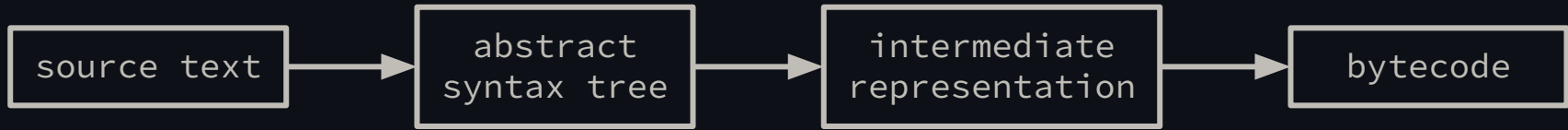




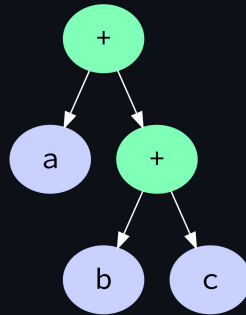
"a + (b + c)"



```
ADD r3, r1(b), r2(c)
ADD r4, r0(a), r3
```



"a + (b + c)"

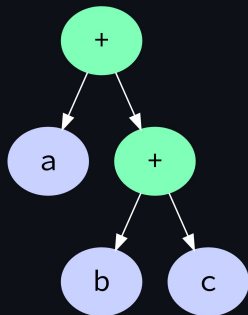


```
ADD r3, r1(b), r2(c)
ADD r4, r0(a), r3
```





"a + (b + c)"

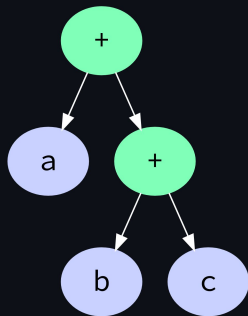


i3 = add b, c  
i4 = add a, i3

ADD r3, r1(b), r2(c)  
ADD r4, r0(a), r3

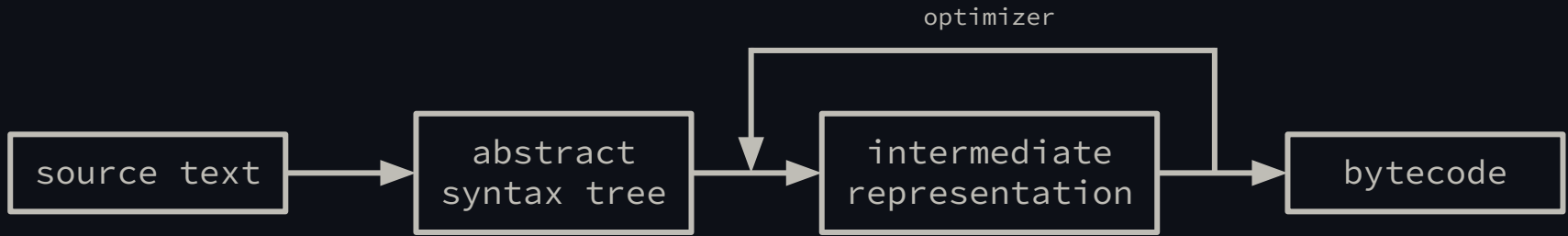


"a + (b + c)"

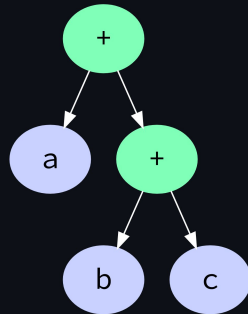


```
i0 = a
i1 = b
i2 = c
i3 = add i1, i2
i4 = add i0, i3
```

```
ADD r3, r1(b), r2(c)
ADD r4, r0(a), r3
```

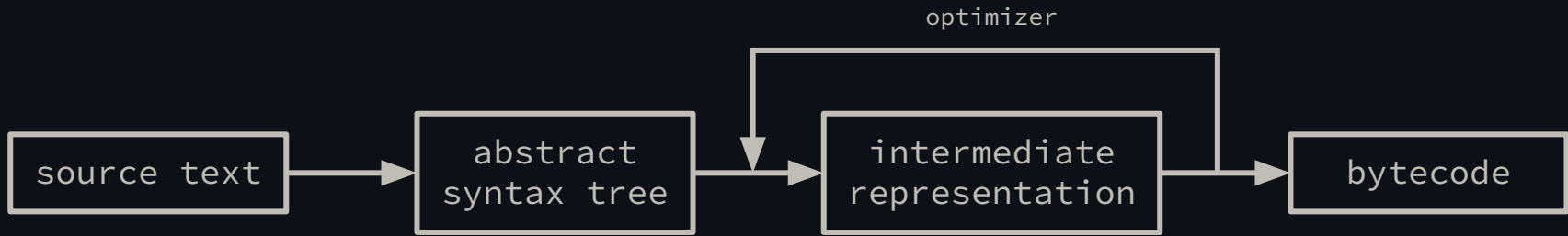


"a + (b + c)"

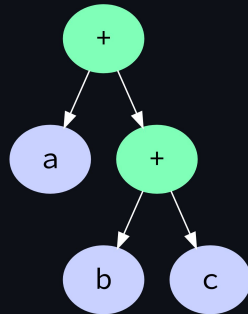


```
i0 = a
i1 = b
i2 = c
i3 = add i1, i2
i4 = add i0, i3
```

```
ADD r3, r1(b), r2(c)
ADD r4, r0(a), r3
```

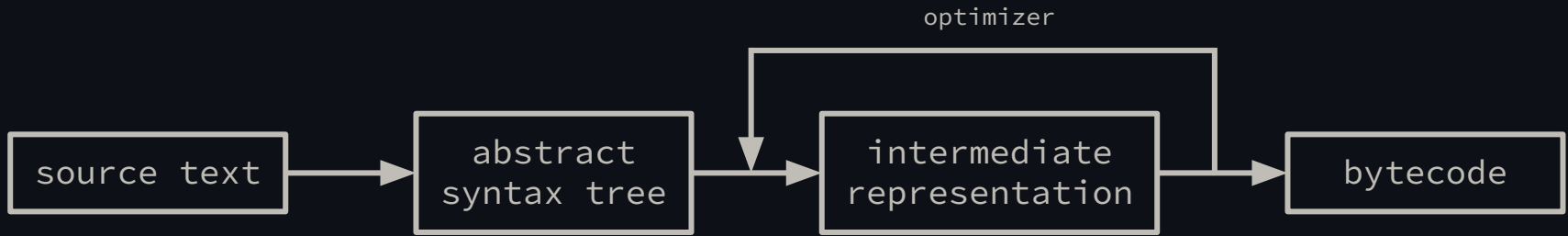


"a + (b + c)"

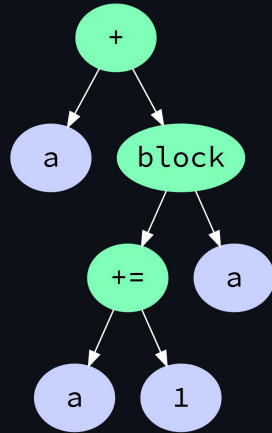


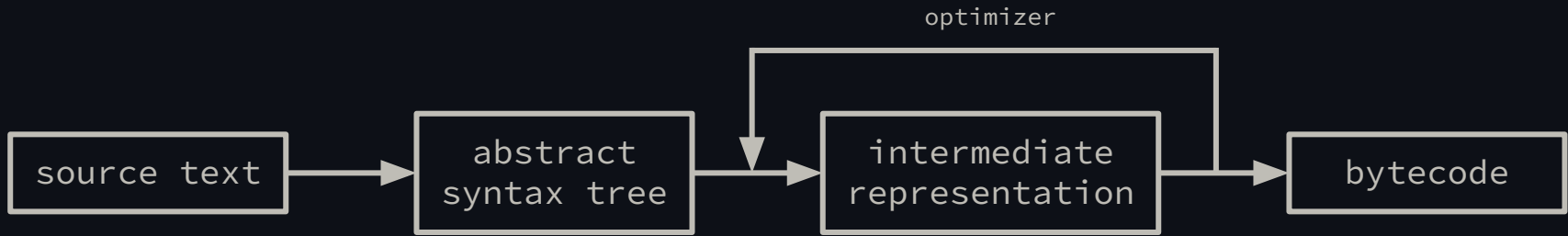
i3 = add b, c  
i4 = add a, i3

ADD r3, r1(b), r2(c)  
ADD r4, r0(a), r3

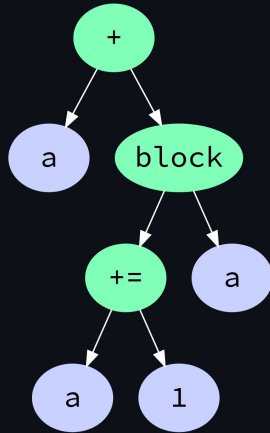


"a + (b + c)"

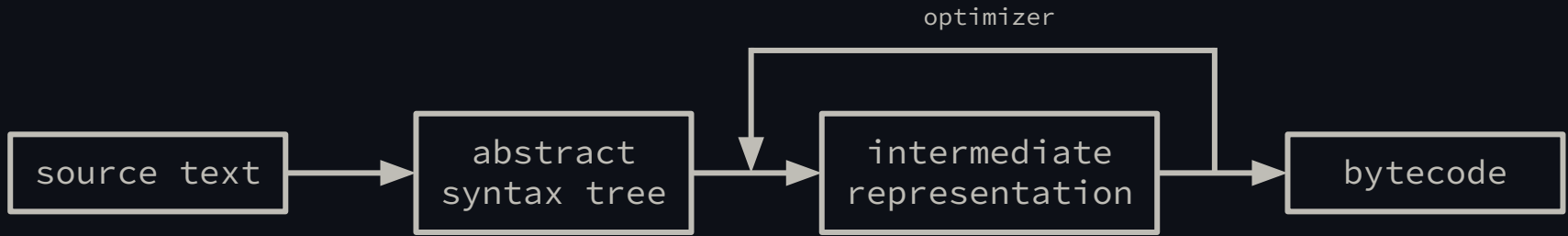




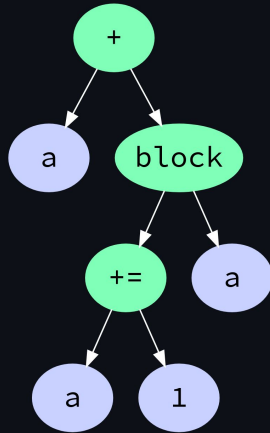
"a + (b + c)"



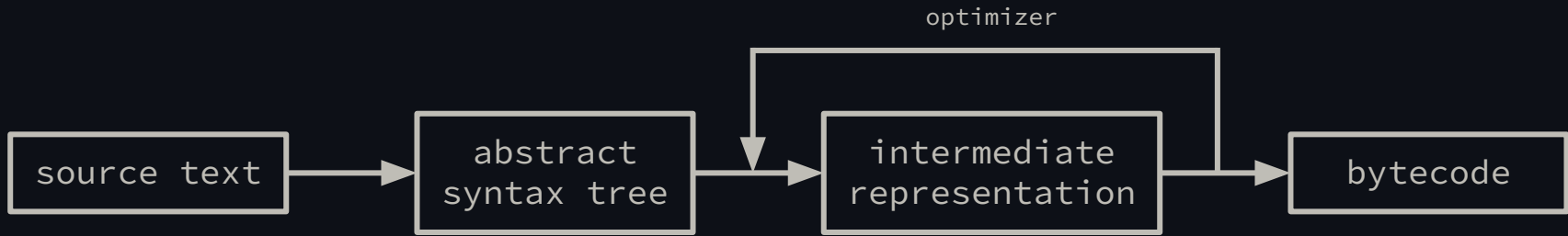
```
i0 = a  
a = add a, #1  
i1 = a  
i2 = i0 + i1
```



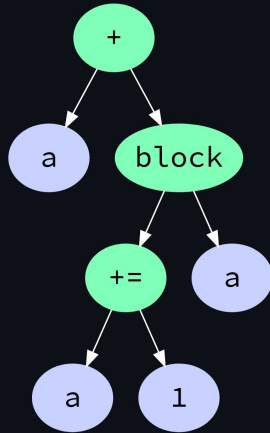
"a + (b + c)"



```
i0 = a
a  = add a, #1
i2 = i0 + a
```



"a + (b + c)"



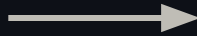
```
i0 = a  
a = add a, #1  
i2 = i0 + a
```

```
COPY r1, r0(a)  
ADD r0(a), r0(a), #1  
ADD r1, r1, r0(a)
```



```
i0 = a  
a  = add a, #1  
i1 = a  
i2 = i0 + i1
```

???



```
i0 = a  
a  = add a, #1  
i2 = i0 + a
```

















```
a + { a += 1; a }
```

```
i0 = a  
a  = add a, #1  
i1 = a  
i2 = add i0, i1
```

```
a + { a += 1; a }
```

```
i0 = a  
a  = add a, #1  
i1 = a  
i2 = add i0, a
```

```
a + { a += 1; a }
```

```
i0 = a  
a  = add a, #1  
i2 = add i0, a
```





```
i0 = a
```

```
...
```

```
if i2 then L7 else L8
```

```
L7: a = add a, #1
```

```
L8: ...
```

```
i5 = add i0, #7
```

# Copy propagation

🌐 4 languages ▾

[Article](#) [Talk](#)

[Read](#) [Edit](#) [View history](#)

From Wikipedia, the free encyclopedia

In [compiler theory](#), **copy propagation** is the process of replacing the occurrences of targets of direct assignments with their values.<sup>[1]</sup> A direct assignment is an instruction of the form `x = y`, which simply assigns the value of `y` to `x`.

From the following code:

```
y = x
z = 3 + y
```

Copy propagation would yield:

```
z = 3 + x
```


Copy propagation often makes use of [reaching definitions](#), [use-def chains](#) and [def-use chains](#) when computing which occurrences of the target may be safely replaced. If all [upwards exposed uses](#) of the target may be safely modified, the assignment operation may be eliminated.


Copy propagation is a useful "clean up" optimization frequently used after other compiler passes have already been run. Some optimizations—such as classical implementations of [elimination of common sub expressions](#)<sup>[1]</sup>—*require* that copy propagation be run afterwards in order to achieve an increase in efficiency.

See also [\[ edit \]](#)

- [Copy elision](#)
- [Constant folding and constant propagation](#)

Discord

 **Handmade Network**

 1 Event

Browse Channels

HELLO WORLD


- welcome
- announcements
- project-announcements
- introduce-yourself


MAIN()

- general
- help 7 New
- fishbowl
- off-topic


VISIBILITY JAM

- jam

 Visual Studio Code

 leddoo Yeet

# language-dev Working on a language? Have questions for language and comp...

 leddoo 02/03/2023 7:38 PM

i'm implementing a new compiler for my scripting language. currently,  $c = a + b$  is compiled to the following IR (it's basically SSA, except that locals are implemented as mutable registers):

```
copy t1, a
copy t2, b
add t3, t1, t2
copy c, t3
```



in the previous version of the compiler, i got rid of those copies by passing a flag into the code-gen function (true if the value was only ever read). for locals, this allowed the code-gen function to just return their (mutable) register, instead of allocating a new one & copying the local's value. but this approach has a few issues:

- 1) it only helps during ast lowering. if optimizations produce similar code, i'm out of luck.
- 2) it's incorrect. eg:  $c = a + \{ a += 1; b \}$ . here, the value of  $a$  does require a copy, cause  $a$  is changed before the actual read occurs (during the `add`).


i was kinda hoping that register allocation would solve this issue. but RA assumes that all live values are actually required. a quick google search suggests that the optimization i'm looking for is called "copy propagation", which "often makes use of reaching definitions, use-def chains and def-use chains".

before going deeper down that rabbit hole, i just wanted to ask, how have you guys approached this problem?







ADMIN — 2

-  AsafGartner
-  cloin Wat


FOUNDER — 1

-  Jeroen


PROJECT CREATOR — 6

-  DanZaidan
-  gingerBill
-  Phillip Trudeau-Tavara
-  raysan5
-  Wassimulator Your Ad Here
-  Zak

PODCAST GUEST — 1





-  vurtun

MEMBER — 96

-  .bmp I see you


Search


Message #language-dev



Discord

**Handmade Network**



1 Event

Browse Channels

HELLO WORLD

- welcome
- announcements
- project-announcements
- introduce-yourself


MAIN()

- general
- help 7 New
- fishbowl
- off-topic

VISIBILITY JAM

- jam

Visual Studio Code

 leddoo  
Yeet


# language-dev Working on a language? Have questions for language and comp...

locals are implemented as mutable registers):



```
copy t1, a
copy t2, b
add t3, t1, t2
copy c, t3
```

in the previous version of the compiler, i got rid of those copies by passing a flag into the code-gen function (true if the value was only ever read).  
for locals, this allowed the code-gen function to just return their (mutable) register, instead of allocating a new one & copying the local's value.  
but this approach has a few issues:  
1) it only helps during ast lowering. if optimizations produce similar code, i'm out of luck.  
2) it's incorrect. eg: `c = a + { a += 1; b }`. here, the value of `a` does require a copy, cause `a` is changed before the actual read occurs (during the `add`).


i was kinda hoping that register allocation would solve this issue. but RA assumes that all live values are actually required.  
a quick google search suggests that the optimization i'm looking for is called "copy propagation", which "often makes use of reaching definitions, use-def chains and def-use chains".  
before going deeper down that rabbit hole, i just wanted to ask, how have you guys approached this problem?

 ratchettfreak 02/03/2023 7:52 PM  
SSA solves that issue by making all registers read-only






ADMIN — 2

-  AsafGartner
-  cloin  
Wat



FOUNDER — 1

-  Jeroen


PROJECT CREATOR — 5

-  DanZaidan
-  Phillip Trudeau-Tavara
-  raysan5
-  Wassimulator  
Your Ad Here
-  Zak

PODCAST GUEST — 2

-  Rudy
-  vurtun

MEMBER — 99

-  .bmp  
I see you

Search

Message #language-dev









# Static single-assignment form

🌐 7 languages ▾

Article [Talk](#)

Read [Edit](#) [View history](#)

From Wikipedia, the free encyclopedia

In [compiler](#) design, **static single assignment form** (often abbreviated as **SSA form** or simply **SSA**) is a property of an [intermediate representation](#) (IR) that requires each [variable](#) to be [assigned](#) exactly once and defined before it is used. Existing variables in the original IR are split into *versions*, new variables typically indicated by the original name with a subscript in textbooks, so that every definition gets its own version. In SSA form, [use-def chains](#) are explicit and each contains a single element.

SSA was proposed by [Barry K. Rosen](#), [Mark N. Wegman](#), and [F. Kenneth Zadeck](#) in 1988.<sup>[1]</sup> [Ron Cytron](#), [Jeanne Ferrante](#) and the previous three researchers at [IBM](#) developed an algorithm that can compute the SSA form efficiently.<sup>[2]</sup>

One can expect to find SSA in a compiler for [Fortran](#), [C](#), [C++](#),<sup>[3]</sup> or [Java](#) ([Android Runtime](#));<sup>[4][5]</sup> whereas in [functional language](#) compilers, such as those for [Scheme](#) and [ML](#), [continuation-passing style](#) (CPS) is generally used. SSA is formally equivalent to a well-behaved subset of CPS excluding non-local control flow, which does not occur when CPS is used as intermediate representation.<sup>[3]</sup> So optimizations and transformations formulated in terms of one immediately apply to the other.

## Benefits [\[ edit \]](#)

The primary usefulness of SSA comes from how it simultaneously simplifies and improves the results of a variety of [compiler optimizations](#), by simplifying the properties of variables. For example, consider this piece of code:

```
y := 1
y := 2
x := y
```

Humans can see that the first assignment is not necessary, and that the value of `y` being used in the third line comes from the second assignment of `y`. A program would have to perform [reaching definition analysis](#) to determine this. But if the program is in SSA form, both of these are immediate:

```
y1 := 1
y2 := 2
x1 := y2
```









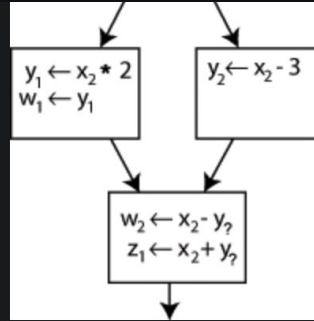












It is clear which definition each use is referring to, except for one case: both uses of  $y$  in the bottom block could be referring to either  $y_1$  or  $y_2$ , depending on which path the control flow took.

To resolve this, a special statement is inserted in the last block, called a  $\Phi$  (Phi) function. This statement will generate a new definition of  $y$  called  $y_3$  by "choosing" either  $y_1$  or  $y_2$ , depending on the control flow in the past.

