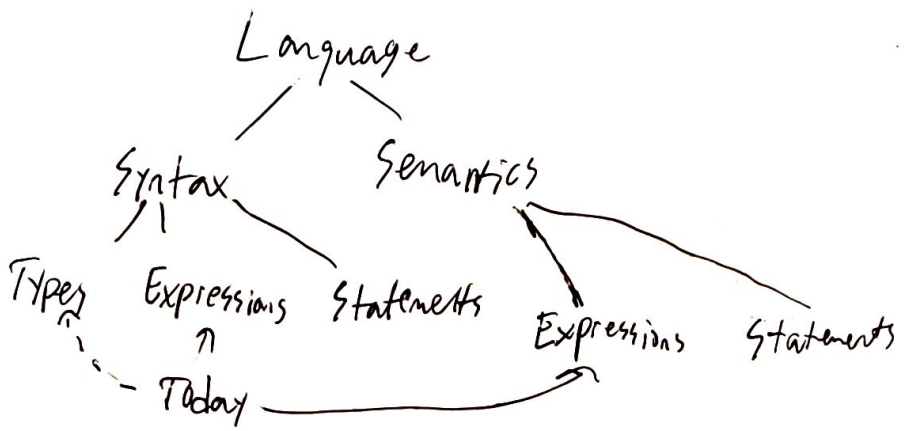
A simple programming language

1/26

Use models of PLs - easier to reason about
More complex features usually aren't that theoretically interesting



Datatypes: Integers, Booleans
others like float, char, string, easy enough to add
Arrays of ints and bools, w/integer indices

BNF (Backus-Naur Form)

e ::= n | true | false | x | e iop e | e ? e : e | e bop e | a[e] | size(a)

expression can be
also y, z, x', ...
also b, a', ...
n: any int
x: variables
iop: +, -, ≤, =, ... min, max, ...
Conditional (C, Java)
bop: ∧, ∨, →, ...
a[e]: array access

Don't have: Assignments, pointers, functions, arrays as values, ...

We'll allow multidimensional arrays. They're syntactic sugar anyway
a[i][j] = a[i * size_x(a) + j]

(x ? a : b)[0] ✗
but (x ? a[0] : b[0]) ok

Examples:

$(x < \bar{0}\ ?\ x+y : x*y)+z$

$(x < \bar{0}\ ?\ \bar{0} : sqrt(x))$

$(i < \bar{0}\ ?\ a[\bar{0}] : i \geq size(a)\ ?\ a[size(a)-\bar{1}] : a[i])$

Functions not values:

$(x > \bar{1}\ ?\ min : max)(y,z)$ ✗

but $(x > \bar{1}\ ?\ min(y,z) : max(y,z))$ OK

Values of Expressions (Semantics)

Syntactic vs. Semantic

What's the value of $\bar{2}+\bar{2}$.? $\bar{4}$

How did we know?

In general, what's the value of $\bar{n}_1 + \bar{n}_2$? $\overline{n_1+n_2}$

$\bar{n}_1$, first +, $\bar{n}_2$ are syntactic (in the program)

n_1, second +, n_2 are semantic (normal math meanings)

That's why I use the $\bar{}$ for syntactic #s.

I'll forget, you'll forget, but it'll usually be clear from context.

Syntactic: true, false

Semantic: $\bar{T}$, $\bar{F}$

Depends on context: $\wedge$, $\vee$, ...

What's the value of $x+y$? Depends on x and y

Need a state σ

$\sigma(e)$ will be the value of e in the state σ

σ must be proper! (Still: assigns a value to all (free) vars)

$\sigma(\bar{n}) = n$

$\sigma(\text{true}) = T$

$\sigma(\text{false}) = F$

Values are always semantic values

$\sigma(x) =$ the value of x in σ

$\sigma(e_1 + e_2) = \sigma(e_1) + \sigma(e_2)$ (same for $-, *, =, \leq, \ldots$)

(syntactic $+$ on the left, semantic $+$ on the right)

$\sigma(e_1 \wedge e_2) = \sigma(e_1) \wedge \sigma(e_2)$ (same for $\vee, \rightarrow, \leftrightarrow, \ldots$)

$$\sigma(e ? e_1 : e_2) = \begin{cases} \sigma(e_1) & \sigma(e) = T \\ \sigma(e_2) & \sigma(e) = F \end{cases}$$

$\sigma(a[e]) = \sigma(a[\sigma(e)])$

$\sigma(\text{size}(a)) = |\sigma(a)|$

$\sigma(x < \bar{0}) = \sigma(x) < \sigma(\bar{0}) = 1 < 0 = F$

$\sigma\{x=1, y=2, z=3\}((x < \bar{0} ? x + y : x * y) + z) = \sigma(x*y) + \sigma(z)$

$= \sigma(x) * \sigma(y) + \sigma(z)$

$= 1 * 2 + 3$

$= 5$

How do we write array values in states?

Semantic value of arrays

A few options: A finite sequence ← We'll use

A set of ordered pairs (relation)

A function from integers to values

$\sigma = \{a = [8; 2; 5], x = 3, y = 0\}$

$\sigma(a[x*y]) = \sigma(a[\sigma(x*y)]) = \sigma(a[3*0]) = \sigma(a[0]) = 8$

$\sigma = \{x = 1; a = [2; 0; 4]\}$

$\sigma(a[x+1] - \bar{2}) = \sigma(a[x+1]) - \sigma(\bar{2}) = \sigma(a[\sigma(x+1)]) - 2$

$= \sigma(a[\sigma(x) + \sigma(1)]) - 2$

$= \sigma(a[1+1]) - 2 = \sigma(a[2]) - 2$

$= 4 - 2 = 2$