SI 413: Some people, when confronted with a problem, think “I know, I’ll use regular expressions.” Now they have two problems

Professor Keith Sullivan

Semantic Analysis

- Semantics concerns a program’s meaning.
- Static and dynamic semantics
  - Static: computed or known at compile time
  - Dynamic: computed or known at run-time
- Huge variance in semantic rules
- Semantic analyzer enforces static rules and annotates program with information for the intermediate code generator
  - Clarifications
  - Requirements for dynamic semantic checks

Dynamic Semantic Checks

- Errors are less likely in production, but cost more
- Execution speed?
- Assertions
  - Java: assert denominator != 0;
  - C: assert (denominator != 0);

Static Analysis

- Type checking
- $(7 < 2) + 3$
- Precise
- Eliminate expensive unnecessary dynamic checks
- Statically-typed languages

Attribute Grammars

- Associate meaning with nodes of the parse tree
- Individual rules could have multiple attributes
  - Type, symbol table, intermediate form, list of semantic errors, file name

Example

$$\begin{array}{lcll}
E & \rightarrow & OPA\ E\ T & \triangleright\ E.val = E.val\ OPA.op\ T.val \\
E & \rightarrow & T & \triangleright\ E.val = T.val \\
T & \rightarrow & OPM\ T\ F & \triangleright\ T.val = T.val\ OPM.op\ F.val \\
T & \rightarrow & F & \triangleright\ T.val = F.val \\
F & \rightarrow & -\ F & \triangleright\ F.val = -1 * F.val \\
F & \rightarrow & (\ E\ ) & \triangleright\ F.val = E.val \\
F & \rightarrow & NUM & \triangleright\ F.val = NUM.val
\end{array}$$

Synthesized Attributes: values are calculated (synthesized) only in production rules in which their symbol appears on the l.h.s

Inherited Attributes

- Inherited Attributes are calculated when their symbol is on the r.h.s of the production rule
- Allows information to go down parse tree
  - Symbol table information
  - External environment

Abstract Syntax Trees

- AST are **not** about syntax!
- Simpler than parse trees, since represents only the *meaning* of the program
- Non-unique
- Nodes are either statements or expressions
- Ordering is shown by nesting: the last child of a statement is the next statement

Are ASTs language dependent?

Example

Given this grammar, determine if is is SLR parsable (and fix it if necessary). Then write the AST, with synthetic and inherited attributes, for $x := (5 + 3) * 2; x - 7;$:

$$
\begin{array}{lcl}
run & \rightarrow & stmt\ run \mid stmt \\
stmt & \rightarrow & ares\ STOP \\
ares & \rightarrow & VAR\ ASN\ bres \mid bres \\
bres & \rightarrow & bres\ BOP\ res \mid res \\
res & \rightarrow & res\ COMP\ exp \mid exp \\
exp & \rightarrow & exp\ OPA\ term \mid term \\
term & \rightarrow & term\ OPM\ factor \mid factor \\
factor & \rightarrow & NUM \mid VAR \mid LP\ bres\ RP
\end{array}
$$