Translating Cypher into SIM

Carson Ball
Using a grammar to solve parsing problems

• Given input string, does it belong to language? What does it mean? Ex: Calculator, Compiler
• Grammar: Set of rules for producing strings
• String belongs to language if producible by grammar
Using a grammar to solve parsing problems

• Example:

\[
\begin{align*}
<\text{Exp}> & ::= <\text{Exp}> + <\text{Term}> | <\text{Exp}> - <\text{Term}> | <\text{Term}> \\
<\text{Term}> & ::= <\text{Term}> * <\text{Factor}> | <\text{Term}> / <\text{Factor}> | <\text{Factor}> \\
<\text{Factor}> & ::= x | y | \ldots | ( <\text{Exp}> ) | - <\text{Factor}>
\end{align*}
\]
ANTLR

• Allows you to specify a grammar
• Generates a parser
• Provides parse tree
• Generates interfaces for visiting parse tree
Process for extending translator

1. Choose a subset of Cypher
2. Write unit tests
3. Verify test cases against databases
4. Add subset to ANTLR grammar file
5. Create data structures for translation
6. Create parse tree visitors
7. Visit tree and translate
8. Repeat
Example Cypher statement

• “MATCH (n:Person) RETURN n”
• Should translate to “FROM PERSON RETRIEVE *;”
MATCH (n:Person) RETURN n

clauses
  : match (where)? returnClause #read
  | create #writenode
  | edgeMatch where edgeCreate #writeedge ;