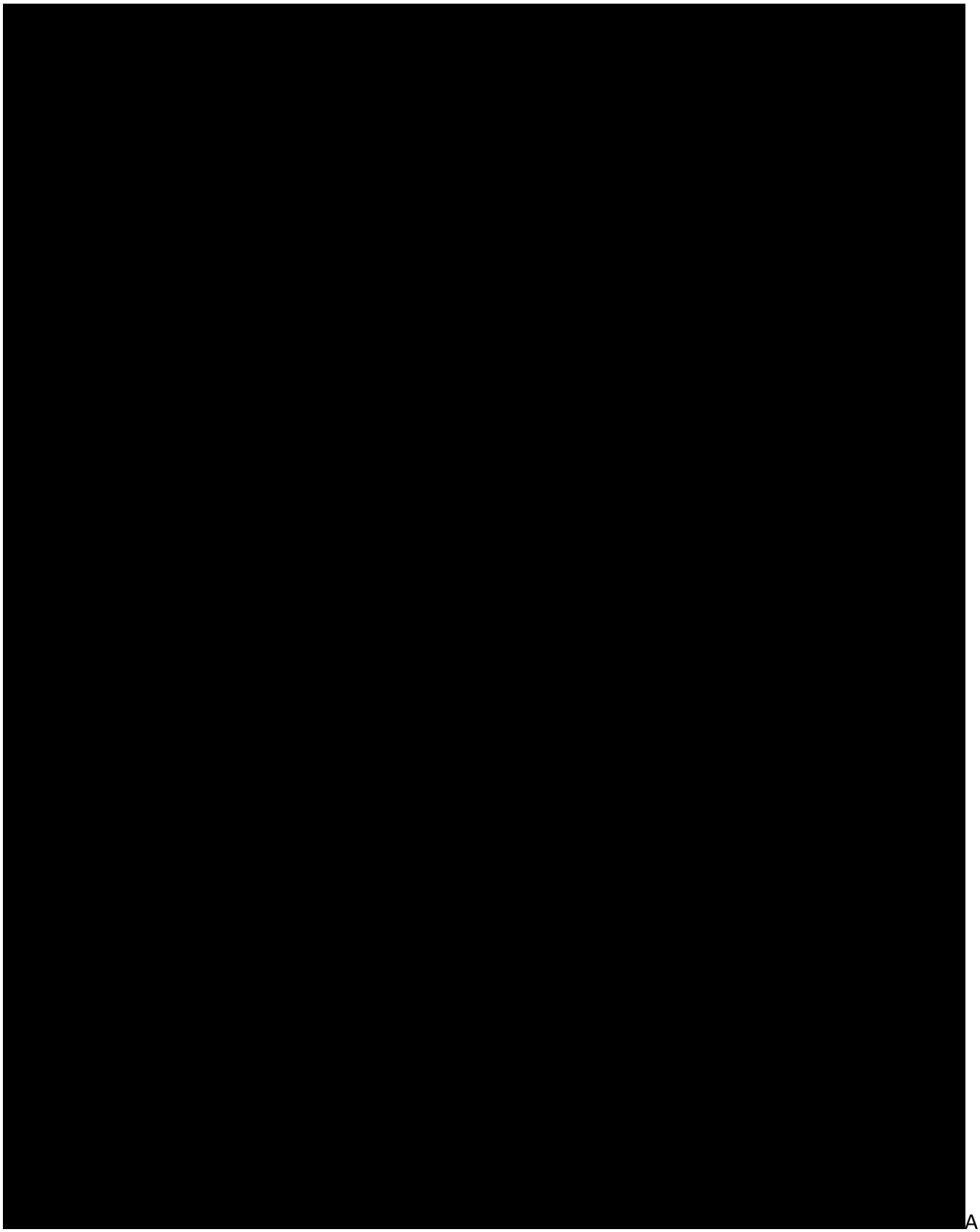
A

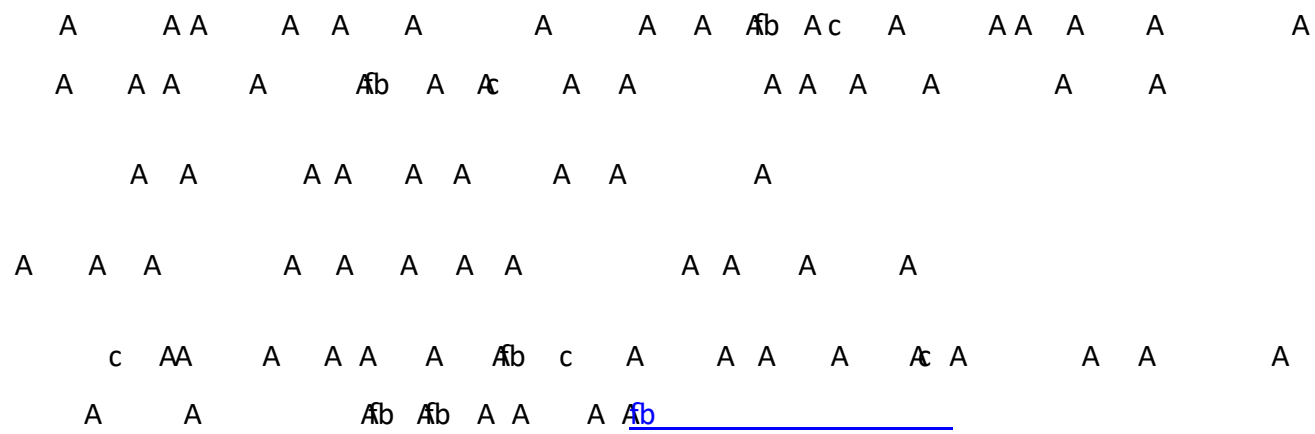


A

A

Learn more at https://tomassetti.me

About this Tutorial



explain the basics

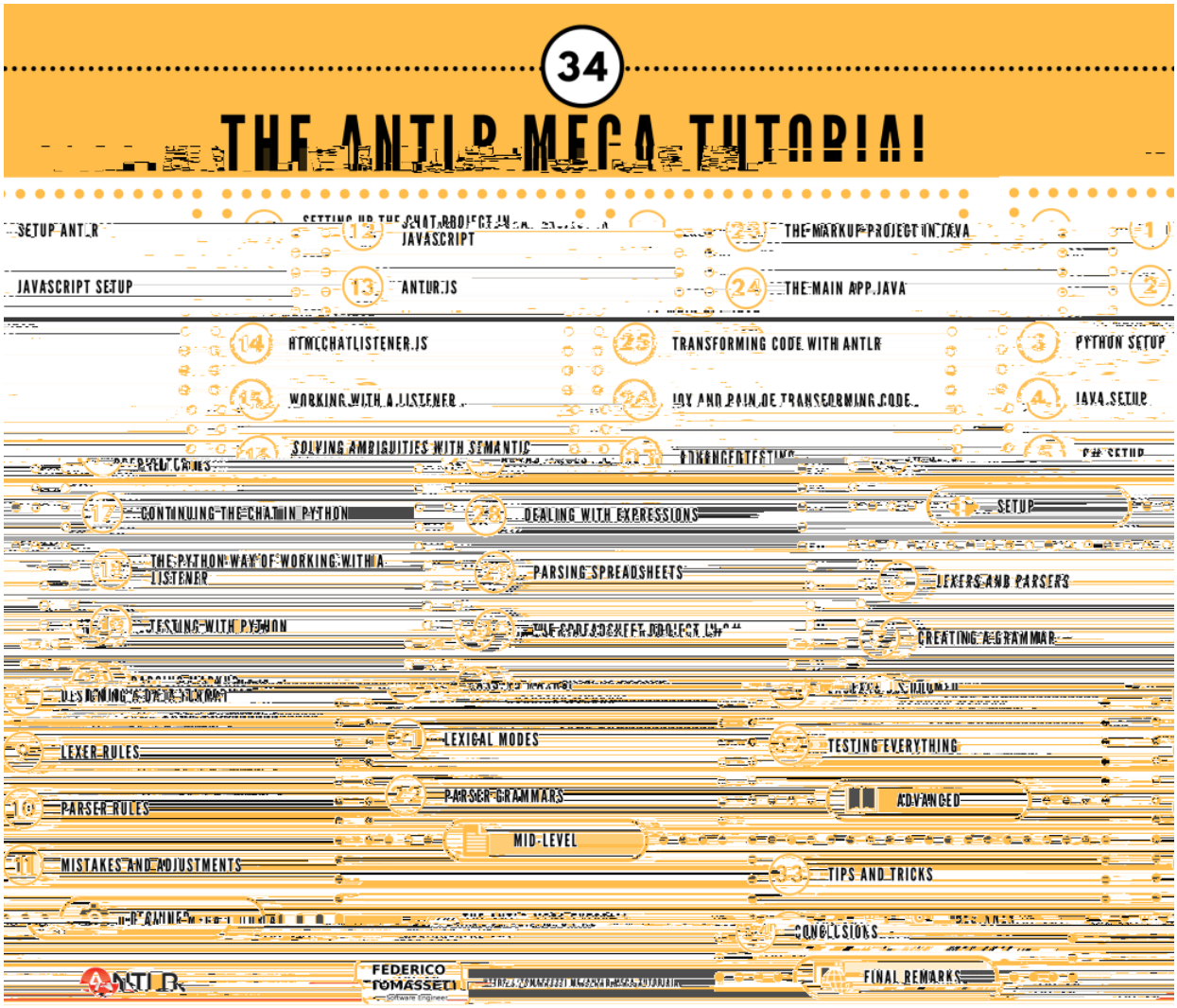
how to setup ANTLR

how to test

advanced and useful features

tons of examples

A



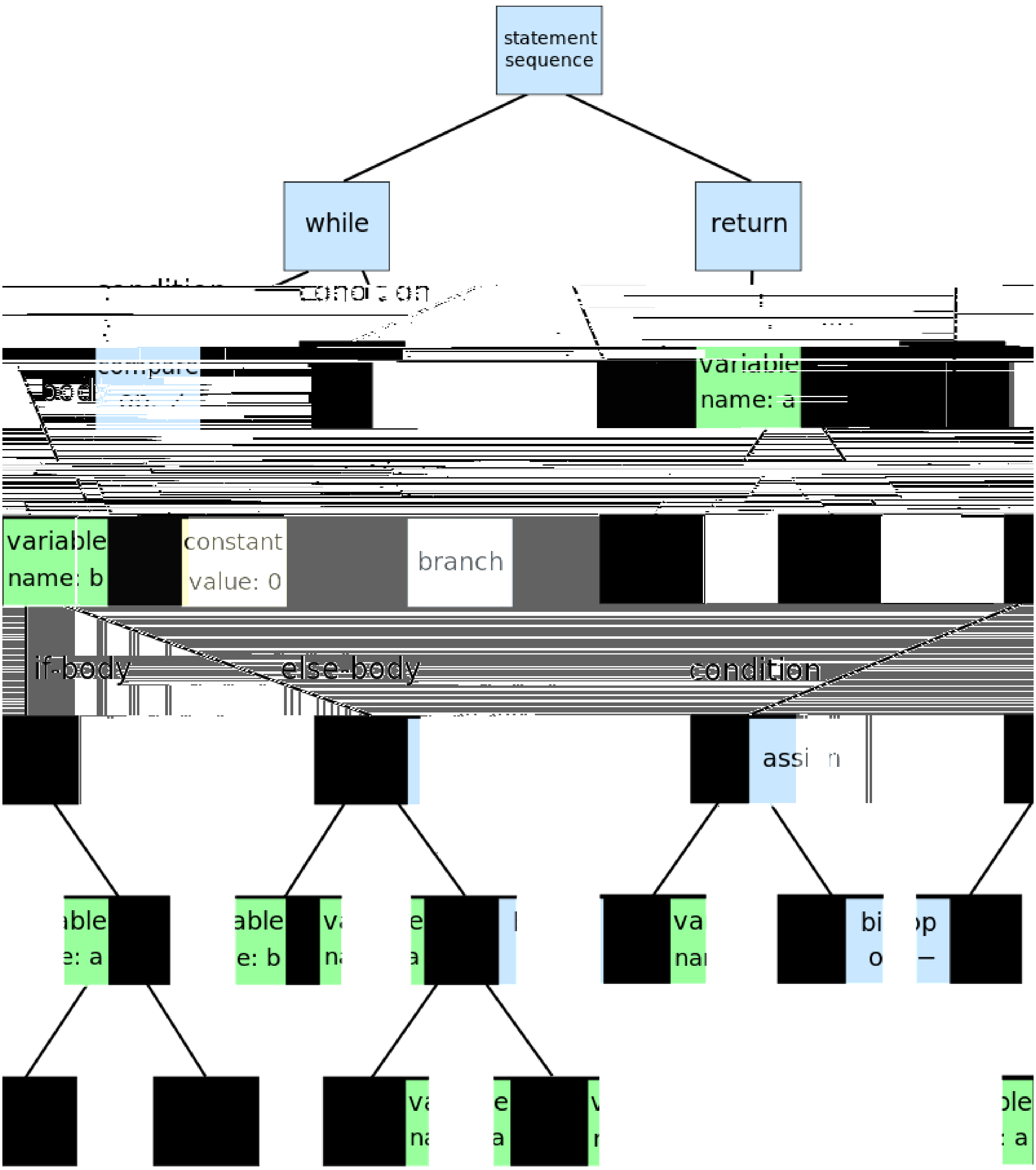
A

AN A A A A fla A

What is ANTLR?

A A A A A A A A A A A A A **A parser takes a piece of text and transforms it into an organized structure** A A *parse tree* A A A A A *Abstract Syntax Tree (AST).* A A A A fla A A A A A A c A A A fla A A A A A A A A A A A A A A A A c A A A A A A A A

A



A

A A fb A A fb A A A A

A A A A A A A A A A A

fb A A A A A A

A A A A A A A A A A A A A A A A A

A A A

A A A A A A A A Ac A A A A A A A

A A A A A A A

Aren’t regular expressions enough?

ANTLR vs writing your own parser by hand

A

Table of Contents

A

A A A

A A A A A fb A A A A A A Afb A A A A A A
A A A A A A A A A
A A Ac A A ffb A Ac A A A Ac A A
c A A A A

Setup

A A A A A A A A A A A A A A
A A A A A A A Afb A A A

1. Setup ANTLR

A A A A A fb A A A A A A A A A A A A A A A
A A A A A A

A A Ac A A Ac A A A A A A A A Ac A A A
Afb A fb A Ac A A

A A A A A A A A A A A A A A A A A A A
A A A A A A A A A A A A A A A A A A
A A ffb Afb A A A A Ac A c Ac A A A A A A A
A A A A A A A A

A A Afb A A A A A A A A A **Java 1.7** A A A A A
A A A A A A A A Afb A A ffb A A A A A A A

A	A

Instructions

A A A A A A A A A Ac A A cA
A AG c A
A A A A A A A A A A A A Ac fb A
A A A A A A A A A fbA A A fb A

Executing the instructions on Linux/Mac OS

A	A A
A	A A A c A
A	A A A A
A	A A A A Ac fb A
A	A c A
A	A fbA A A fb A A A A A A A

A	A A A c A
A	A fbA A A fbA A A A A A A A
A	A A A

Executing the instructions on Windows

A	A A
A	A A A A AG c A A A A fb A
A	A A A A A A A A c A A A fbA A
A	A A A A A Ac A A A A A A A A A
A	A A A A c A A A c A A AW c A
A	A A AG c A
A	A A A A
A	A A c AA
A	A A AA
A	A A c AA
A	A A A
A	A A A A A A A A A fbA A A A A A

A

Typical Workflow

A A A A A Ac A A A*grammar* A Afb A A A A A A A A fbA A A A A A A A A A A A A A A A A Afb A A A A A A A A A A A A A A A

A	A A fb A

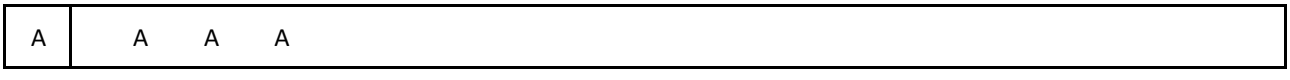
A

A A A A fbA A A A A fbA A A A

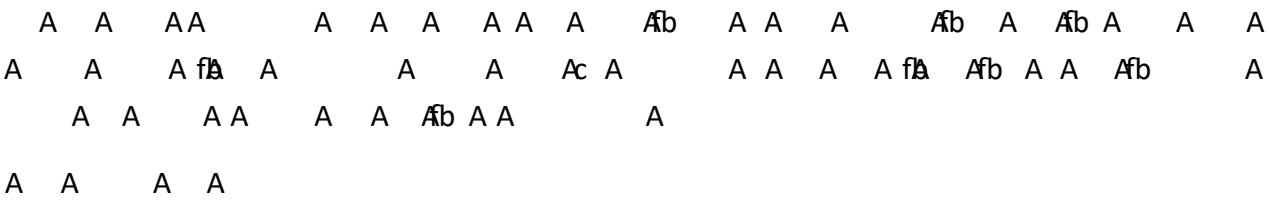
G A A A fbA A A A A A A A A A A A A A A A ffb Afb A A A A A fb A A A A A A A A A A A A A A A AfbA A A A A A A A A A A A A A A A A

C A fb A A A A A A A A A A A A A A A A A A A A A AfbA A A A A A A A A A A A A A A A A A A Ac A A A A fb Ac A

2. Javascript Setup



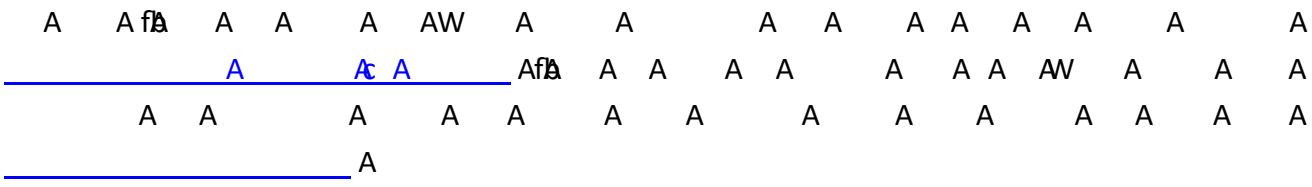
3. Python Setup



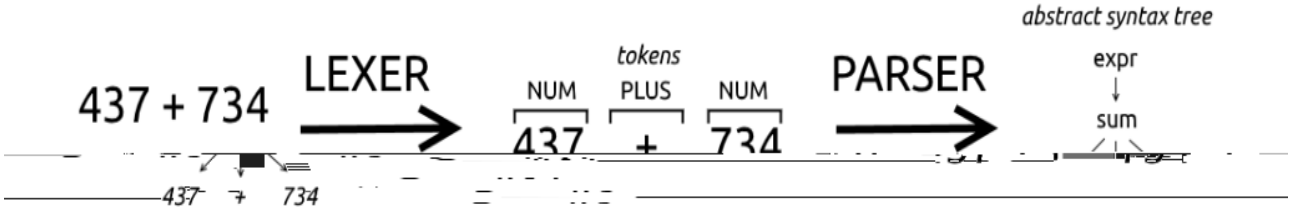
```
A
A
dependencies {A
    antlr "org.antlr:antlr4:4.9.1"A
    compile "org.antlr:antlr4-runtime:4.9.1"A
    testImplementation(platform('org.junit:junit-bom:5.7.0'))A
    testImplementation('org.junit.jupiter:junit-jupiter')A
}A
A
generateGrammarSource {A
    maxHeapSize = "128m"A
    arguments += ['-package', 'me.tomassetti.examples.MarkupParser', '-visitor', '-no-listener']A
}A
compileJava.dependsOn generateGrammarSourceA
A
sourceSets {A
    generated {A
        java.srcDir 'generated-src/antlr/main/'A
    }A
}
compileJava.source sourceSets.generated.java, sourceSets.main.java

clean{
    delete "generated-src"
}

idea {
    module {
        sourceDirs += file("generated-src/antlr/main/")
```

```
    }
}

test {
    useJUnitPlatform()
    testLogging {
        events "passed", "skipped", "failed"
    }
}
```

A A A A A A A Afb A A A A A A Afb A
A A A A A A A A A A A fbA A A A A A A A
A A A A A A Ac A A A



A A A A fb A A c A



A

A A A A AC A A A A A

A	/*A
A	A A A
A	A
A	AA
A	AAA N C A A N C AA
A	AA
A	A
A	A A A
A	A
A	AA
A	N C AA A AA
A	AA
A	AAAA A AA

A A A A A Ac A A A A A A A A A A A A A A
A A A A A A A A Ac A A A A A Afb A A fA
A Ac A A A A A A A Afb A A
A A A A A A A A A A A A A A A A A
A A A A A A A A A A A A c A **Alexer rules are analyzed in the order that they appear** A A A Ac A c A

A A A A A fb A A A A A A Ac A A A fA A
c A A c A A A A A fb A Afb c Ac A A A A*class*A
A A*function* A A A A fA A A A A A c Ac A A Afb A A A A A
A A fb A A A A A*class*A A*function*A A fb Afb A A A Afb A A
fb A A A A A

Ac A A fA A A A **Athere is a name, a colon, the definition of the rule and a terminating semicolon**A

NUMBER

WHITESPACE

7. Creating a Grammar

Top-down approach

Bottom-up approach

A A fb A A Afb A Ac A Ac A A A A A A A A A
A A A Afb A A A A A A A A A A A A A A A
A Afb A AA A A fa A Ac A Afb A A A A A A A
A A A A A A Ac A

A A A A A A AG A A A A A fa A A A
A A A A A A A A A A Afb A A A A A A
A fb A A A A c A A A A A A A Afb A
A A A A N A Afb A A A A A A A Ac A A A fb A
A

A A fa Ac A A A A A Afb A A A A A A A A
A A c A A A A A Ac A A A A A A Ac A A A A
A A A fb Afb A Ac A A A A A A A A A A A A
A A fb Afa A A A A A A A A A fa A A A A

8. Designing a Data Format

A A Afb A A A A A fb A A A A A A A A A
A A A Ac A A c A A A A A c A A Ac A A
A A A A A A A A A fa A A

A A A A A A A A Afb A A A A A

A A A Ac A A fa A c A

- A A A A Ac A A A A A A A A A A Ac A
  A A
- A A A A A A A A A A A A A A A N A A
- A A A A A Ac Afb A A A A A A A A A A A A
  A A Afb A A A fa A A

G A A A A A A A A A A Ac A A

9. Lexer Rules

A A A fb A A Afb A A A A c A A A A A A A
A A fa Afb A

A	/*A
A	A A A A
A	A
A	AA
A	fb A AAAAAAAAA AA
A	fb A AAAAAAAAA AA

A	fb A AA AA
A	fb A AA AA
A	fb A AA AA
A	fb A AA AA
A	fb A AA AA
A	AA
A	fb A AA AA
A	fb A AA AA
A	AA
A	AA A A A A A
A	AA
A	AA A A A A A A
A	AA
A	AA A A AA
A	AA
A	AA A AA AA
A	AA
A	AA A A AA AA
A	AA
A	Y AA A AA

A

A A A A A **A fragments** A A A c Ac Ac Afb A A A A
fb A A A A A fb A A A A A A fb A A fb A Ac A A A A A
A A A A A A A fb fb A

A fb A Afb Afb A A A A A A A A A A A AC A A A
A A A A A A A A A A fb A A fb A
A A A A A A Afb A c A A A A Ac fb A A fb A
A A A A A A Afb A A

A	fb A AA AA
A	N C AA A A A AA

A

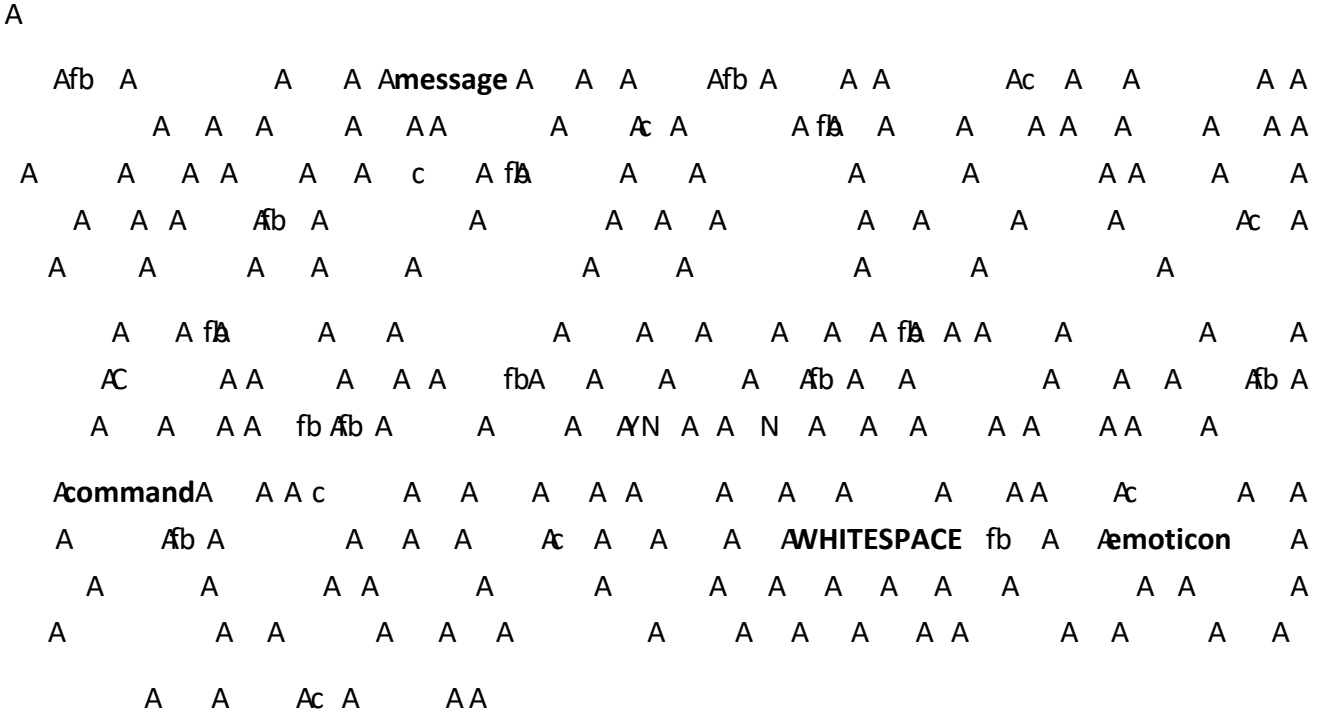
A**TEXT**A A A A A A A Afb A A A Afb A A A
A A A A A A Ac A Ac A A A A A A A A fbA A
A fbA A A fbA A A A A A Ac A fb A A A Ac A A

A A A Afb A A Ac A A A A fbb A A A A
A A A A A A A A A A A A A A A
A AA c A fbA A A

10. Parser Rules

A A A A A A A A A A A A A A A A
A

A	/*A
A	A A A
A	A
A	AA
A	AAAAAAAAAAAAAA A A GAA
A	AA
A	AAAAAAAAAAAAAA A A A A A
A	AA
A	AAAAAAAA A AA AA AA AA AA AA AA
A	AA
A	AAAAAAAAAAAAAA A A A
A	AA
A	AAAAAAAA A AA A A AA
A	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
A	AAAAAAAAAA A A A A
A	AAAAAAAAAAAAAAAA A A A A
A	AAAAAAAAAAAAAAAA
A	AA
A	AAAAAAAAAAAAAA A A Y A A A Y A AA
A	AA
A	AAAAAAAAAA A A A A A A
A	AA
A	AAAAAAAA A A AA


message
command
WHITESPACE
emoticon

A Afb A A A A Afb A A fb A A A ffb A A fb A
A A A A A A A A A A A A A A A A
A A A

A	A A Ac A A A A
A	AA A A A A A
A	AA A Afb A A A
A	A A A
A	A A A
A	A AA A A A AA A A fA A A A A A A A AA A
A	A A A A
A	A A A A A A A A A
A	A A A A A A A
A	A A
A	A A A A A A A A A A A A A A A

A

A A A A A A A A**WORD** A A A A A A Afb A A A A
A A A A A A A A A A A

A	A A A A A
A	A A A A A A A A A
A	A A A A A A A A Y A
A	A G G A

A

A A A A A**TEXT**A AC A A A A A A A fA A A A A A
c A A A A A A A A A A A c A A A A A A A A A**TEXT**A
A A A A A A c A A A A A Afb A fA A A A A
A fA A AC Afb A A A A A A A A A A

A	A
A	AA
A	AAAAAAAAAAAAAAA Y A Y AA
A	AA

A	A
A	AA
A	Y AAAAAAAAAAAAAAAAA A A A

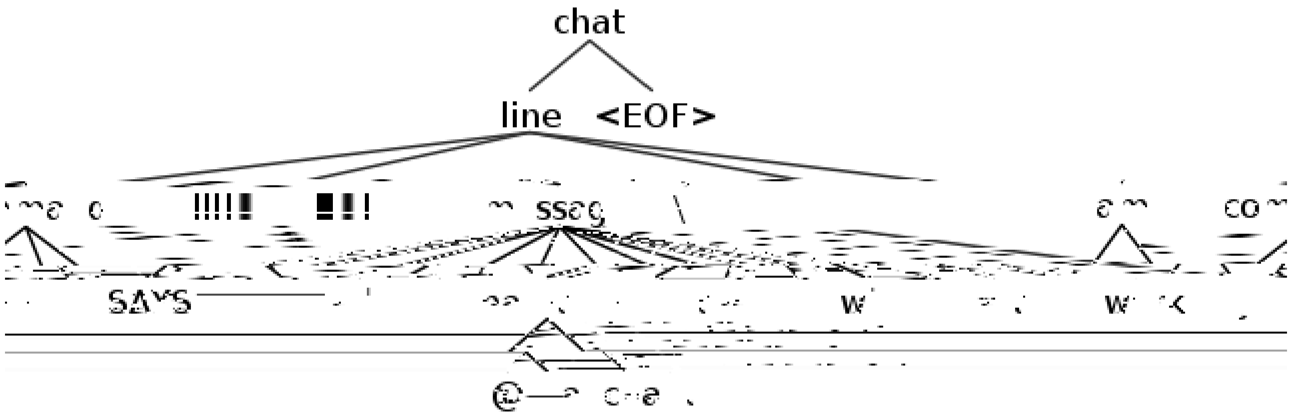
A

A A A A c A A A A A A A A A A A
c A A A A A A A A Ac A A A A A A fbA
A A Ac AC A A AAfb A A A A A A Afb A A

A AfbA A A

A	A A A A A

A A A N AfbAAAA A A A N AAAA A A AfbAA



A

A A A A A A A

A Ac A A A A A A A A AC A A A A A A A A A A
c A A A A A fl A A A A A A A A A A A

A	Y A A A A A

A

A A A Afb A fl A A ATEXT A A A A A A A A A A A
A A Ac A A A A A A A A A A A A A
A A A A A A c A A A A Afb A A A A A
A A A A Ac A

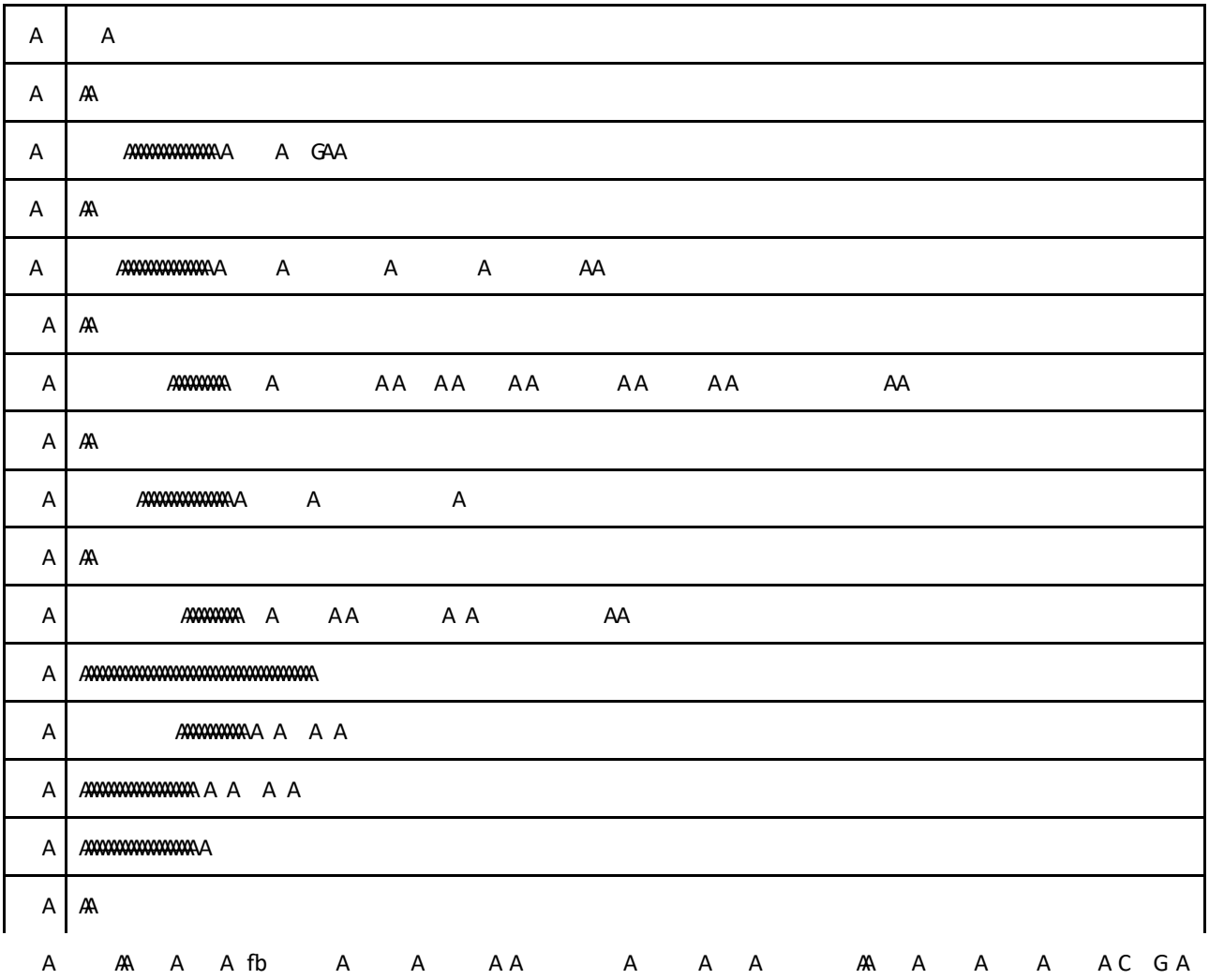
Mid-Level

A A A A A A A A A A A A A Ac A Afb A A A A
A A A A A A A A A A A A A A A A A A A A Ac A
A A A A fl A Ac Ac A A A A A A A A A
A A A A A A A A A A A A A A
c A A A A A A Ac A A A A A A A A c A A
A A A A A

12. Setting Up the Chat Project with JavaScript

A A A A A A A A Ac A A Afb A A A AA Ac A A
A A A A A A A A A A Afb A fl A A Afb A

A	A A
A	AA
A	A
A	A A A



A	fb A AAAAAAAAA AA
A	fb A AAAAAAAAA AA
A	fb A AAAAAAAAA AA
A	AA
A	fb A AAA AA
A	fb A AAA AA
A	AA
A	AAAAAAAAAAAAAA A A A AA
A	AA
A	AAAAAAAAAAAAA A A A A A AA
A	AA
A	AAAAAAAAAAAAAA A A A A AA
A	AA
A	AAAAAAAAAA A A A AA
A	AA
A	AAAAAAAAA A A A A AA
A	AA
A	Y AAAAAAAAAAAAAA A A A

A

A A A A A A A Ac A fb A A A A A A
A A A

A	A A A

A

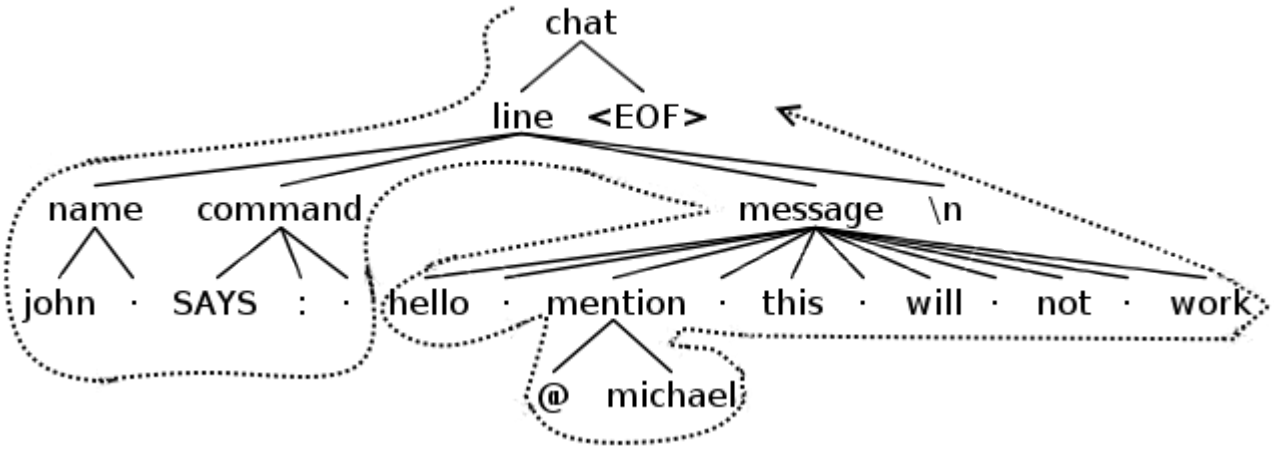
A A Afb A A Afb A A Afb A A A A A A A A
A A A Afb A A fbA A A A Afb A A A A A
A A A A A fbA A

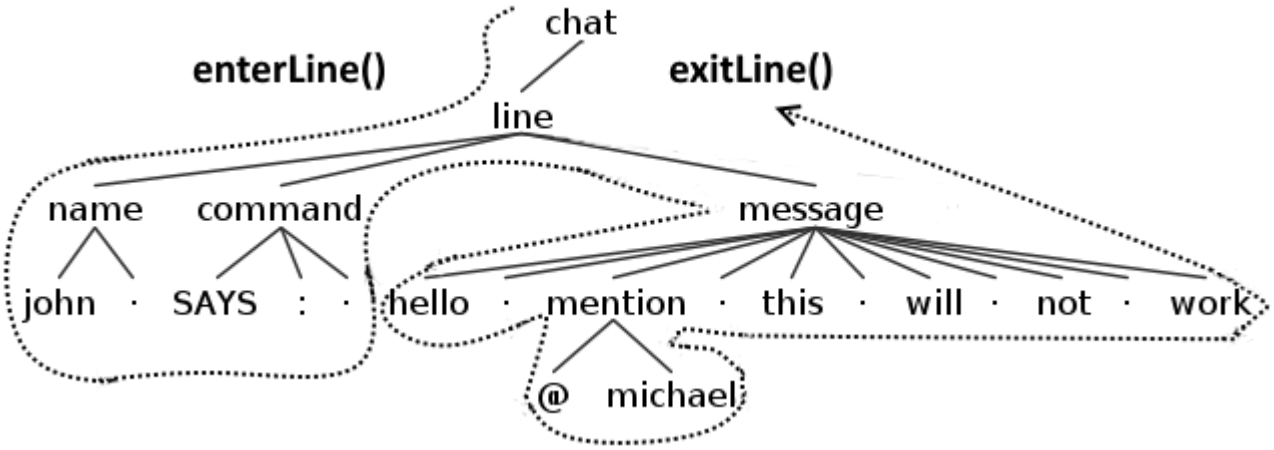
Afb A A A A A A A AA A A A A A fbA A A Ac A A
A A Afb A A A A A A A A A A A A A A
fbA Ac A A Ac A A A A A A A A

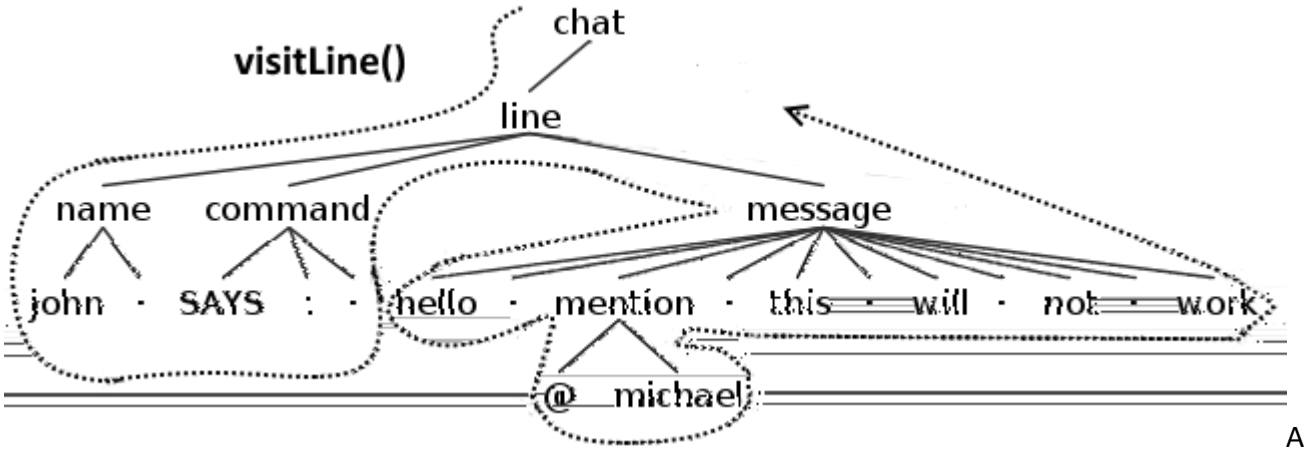
A A A A A A A Afb A A Afb A A fbA A A A A
fb A Ac A A A A A fbA A A A A Ac A A A A A
fb A A fbA A A A A A A A A Afb A A A
A A A A A A A A A A A

A	`import antlr4 from 'antlr4';A`
A	A
A	`// This class defines a complete listener for a parse tree produced by ChatParser.A`
A	`export default class ChatListener extends antlr4.tree.ParseTreeListener {A`
A	A
A	`// Enter a parse tree produced by ChatParser#chat.A`
A	`enterChat(ctx) {A`
A	`}A`
A	A
A	`// Exit a parse tree produced by ChatParser#chat.A`

A exit.


chat
line
<EOF>
name
command
message
\n
john
·
SAYS
:
·
hello
·
mention
·
this
·
will
·
not
·
work
@
michael


chat
enterLine()
exitLine()
line
name
command
message
john
·
SAYS
:
·
hello
·
mention
·
this
·
will
·
not
·
work
@
michael


chat
visitLine()
line
name
command
message
john
·
SAYS
:
·
hello
·
mention
·
this
·
will
·
not
·
work
@
michael

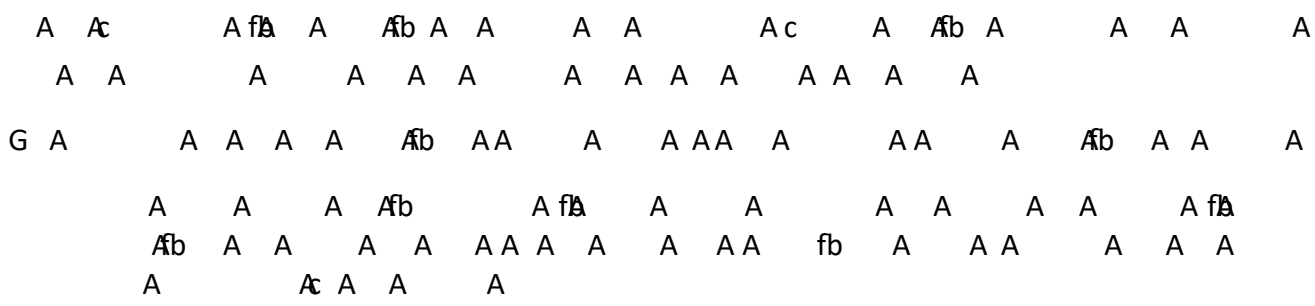
c A Athis is true for the default implementation of a visitor and it's done by returning the children of each node in every function A fiA A A A A fiA A A A A
c A A A A A A A A A A A A

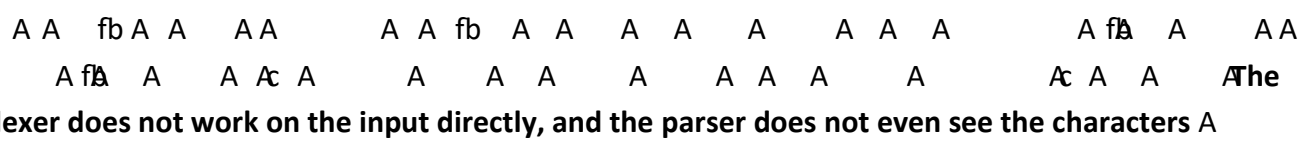
13. Antlr.js

A Afib A A A A A A A A A A

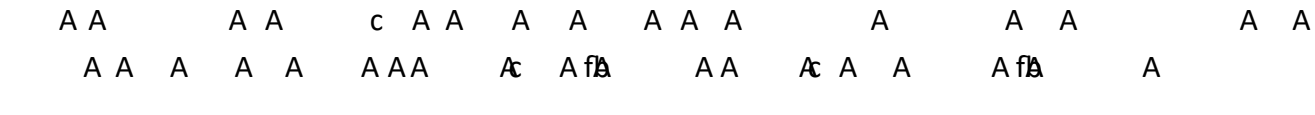
```
import { createServer } from 'http';A
import antlr4 from 'antlr4';A
const { CommonTokenStream, InputStream } = antlr4;A
import ChatLexer from './ChatLexer.js';A
import ChatParser from './ChatParser.js';A
import HtmlChatListener from './HtmlChatListener.js';A
A
createServer((req, res) => {A
  A
  res.writeHead(200, {A
      'Content-Type': 'text/html',        A
  });A
A
  res.write('<html><head><meta charset="UTF-8"/></head><body>');A
  A
  var input = "john SHOUTS: hello @michael /pink/this will work/ :-) \n";A
  var chars = new InputStream(input, true)   A
  var lexer = new ChatLexer(chars);A
  var tokens  = new CommonTokenStream(lexer);A
  var parser = new ChatParser(tokens);A
      A
  parser.buildParseTrees = true;   A
  var tree = parser.chat();   A
  var htmlChat = new HtmlChatListener(res);A
  antlr4.tree.ParseTreeWalker.DEFAULT.walk(htmlChat, tree);A
```

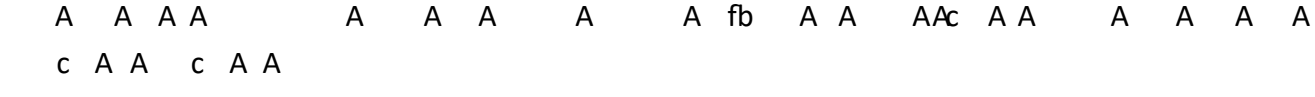
```
  
  res.write('</body></html>');
  res.end();

});
```

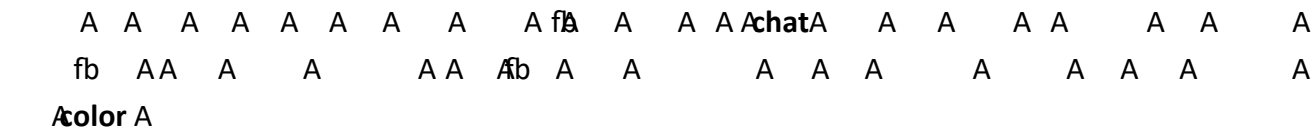




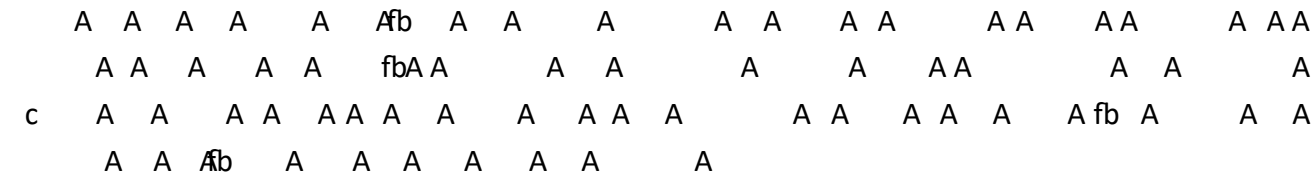
The lexer does not work on the input directly, and the parser does not even see the characters







chat **color**



14. HtmlChatListener.js

Chat

```
import antlr4 from 'antlr4';
import ChatLexer from './ChatLexer.js';
import ChatParser from './ChatParser.js';
import ChatListener  from './ChatListener.js';

```

```
A export default class HtmlChatListener extends ChatListener {    A
A     constructor(res) {A
A         super();A
A         this.Res = res;A
A     }A
A A
A     enterName(ctx) {          A
A         this.Res.write("<strong>");    A
A     }A
A     A
A     exitName(ctx) {      A
A         this.Res.write(ctx.WORD().getText());A
A         this.Res.write("</strong> ");A
A     }A
A         A
A     exitEmoticon(ctx) {      A
A         var emoticon = ctx.getText();        A
A         A
A         if(emoticon == ':-)' || emoticon == ':)')A
A         {        A
A             ctx.text = "☺";A
A         }A
A         A
A         if(emoticon == ':-(' || emoticon == ':(')A
A         {          A
A             ctx.text = "☹";A
A         }A
A     }   A
A     A
A     enterCommand(ctx) {          A
```

```
        if(ctx.SAYS() != null)
            this.Res.write(ctx.SAYS().getText() + ':' + '<p>');

        if(ctx.SHOUTS() != null)
            this.Res.write(ctx.SHOUTS().getText() + ':' + '<p style="text-transform: uppercase">');
    }
    exitLine(ctx) {
        this.Res.write("</p>");
    }
}
```

fb A A Afb A A A A A**HtmlChatListener**A A A**ChatListener.** A A fb A A A A A A

A **ctx**A A A A A fb A fb A A A fb A A A A A A A A fb A A A fb A A A A A A A fb A A A A A A Afb A A A A A A c A A A A A A A A c A G A A A A Afb A A**WORD()**A A**WHITESPACE();**A A A Afb A A**WHITESPACE()** A**SAYS()**A A**SHOUTS().**A

A fb A A A A A A c A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A c A A A A c A **you can't stop the execution of the walker and the execution of the functions** A

A A A A Ac A A A A Ac A A A A A Ac Ac A A A A A A A Afb A A A**WORD**A A A A A A A A A A **WHITESPACE**A A A A A A A A A A A A A A A A A A A A A A A A Afb A

A Afb A*exitEmoticon*A A A fb A A A A A A A A A A A A A fb A A A Ac A A A A A fb Afb A A A A A *enterCommand* A A A Ac A A fb A A**SAYS**A A**SHOUTS** A A A A A A A fb A A A A A Afb A Ac A fb A A Afb A A**SHOUT.** A A A A A A A A A A fb A**line**A Ac A A A A A A A A A fb A A

A A A A A A A A A A A A A Ac A A A
A A A A A A Ac A A A Afb A
A

john SHOUTS:

☺

A A A A A A A A A A A ffb A A A A A A fb A A
A A AcolorA Amessage A

15. Working with a Listener

A A A A A A A fb A A A A A A A A A A A A A A A
A c A A A Ac A A Afb AcolorA A A A A fb A
A A

```
enterColor(ctx) {   A
    var color = ctx.WORD().getText();          A
    this.Res.write('<span style="color: ' + color + '">');          A
}A
A
exitColor(ctx) {          A
    this.Res.write("</span>");     A
}A
A
exitMessage(ctx) {               A
    this.Res.write(ctx.getText());A
}A
```

A

john SHOUTS:

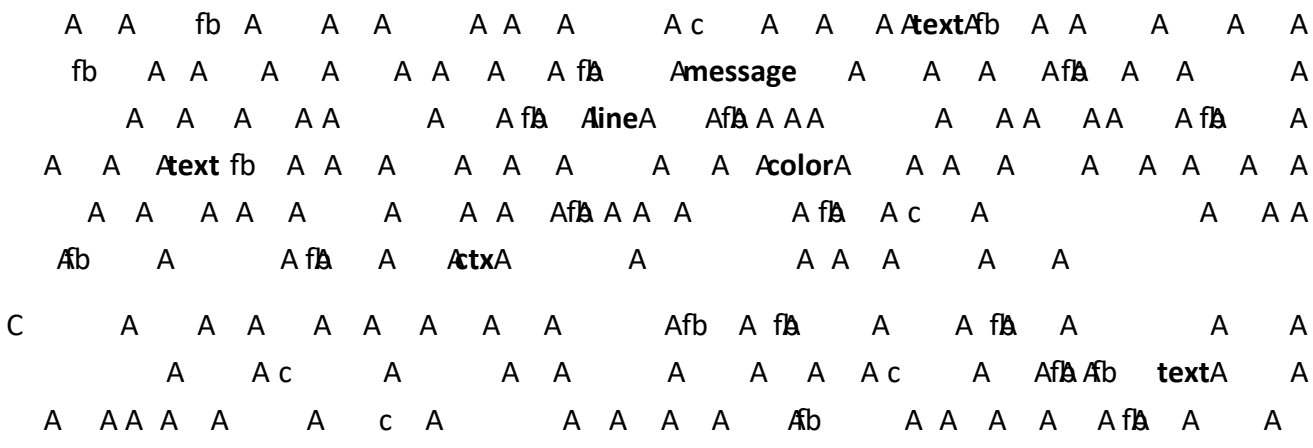
THIS WILL WORK@HELLO @MICHAEL (PINK(THIS WILL WORK

A A A A A A A c A A A A A A A A A A fb message A
A A A A fb A A A A A fb A A A fb message A A A A
A A

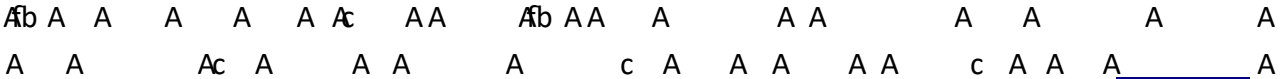
A A A A A A A c A A A A A A A fb A fb A A
A A A c A A

A	`exitColor(ctx) {` A
A	`    ctx.text += ctx.message().text;` A
A	`    ctx.text += '</span>';` A
A	`}` A
A	A
A	`exitEmoticon(ctx) {` A
A	`    var emoticon = ctx.getText();` A
A	A
A	`    if(emoticon == ':-)' || emoticon == ':)')` A
A	`    {` A
A	`        ctx.text = "☺";` A
A	`    }` A
A	A
A	`    if(emoticon == ':-(' || emoticon == ':(')` A
A	`    {` A
A	`        ctx.text = "☹";` A
A	`    }` A
A	`}` A
A	A
A	`exitMessage(ctx) {` A

```
    var text = '';
A
    for (var index = 0; index <  ctx.children.length; index++ ) {
        if(ctx.children[index].text != null)
            text += ctx.children[index].text;
        else
            text += ctx.children[index].getText();
    }
A
    if(ctx.parentCtx instanceof ChatParser.LineContext == false)
    {
        ctx.text = text;
    }
    else
    {
        this.Res.write(text);
        this.Res.write("</p>");
    }
}
```



16. Solving Ambiguities with Semantic Predicates



C fb A A A A A A A A c A ATEXTA A A A A A A
c A Afb Afla A A A A A A Afla A A A A A A A A A
A A Ac A A A A A A

A A A A Afb A fla A *semantic predicates.*A A A A A A
A A A AAc A A A A c A A c A Afb A A
A A A A A A c A A A A A Ac A A A A A
ffb A A fla A A Afb A A A A A A A A A A A
A A

A A A A fla A A A fla*actions* A A A A c A c A A A
A A**The downside is that the grammar is no more language independent** A A A
A A A A A Ac A Afb A A A AG A A A A A A A A
A A A A A A A A A Ac A A A A A fla A
A A A A

A	AAAAAAAAAAAA A A Y A A A Y A A
A	AA
A	Y AAAAAAAAAAAA fb A A A A fb A A A AA

A

A AlinkA A A Afb Ac A A A A A A A A ATEXTA A
A A Ac A Afb Ac A A A A A A A A A A
A A A Ac fb A A A Afla A A A A Ac A A A A
A A A ATEXTA A A A A A A A A Ac A A A A
A A A A A A A A A A A A A A A A A A
A

A A A Afb A A A fbAc A Ac A ffb A A A A ffb A
fb A A Afb A A A A A A A A A A A A A A A A A
A

A A A A Afb A A A A

A	A A A AA A A A A
A	Y AA A A A A A A A AA
A	A A
A	Y AA A A A A A A A AA
A	A A
A	Y AA A A A A A A A AA

fb A A A Afb A A A A A A A A A Ac A A A A
A Afb A A AG A Afb A A A A c A **mention**A A Afb Ac A A
WHITESPACEA A

A	A A
A	A A A A A AA
A	A A
A	A A A A AA
A	A A
A	A fb A AA A A AA
A	A A
A	A A AA A A AA

17. Continuing the Chat in Python

C fb A A A A A A A A fbA A A A A A **TEXT**A Ac fb A
A**WORD**A A A A A A A A A A A A A A A A A
c A A A c A A A A fb A **WORD** A fb A A fb A A A
A

A A A A A A A fb A A A A A fb A A A A A A A
A A A A fb A A A A A

A	A A A

A

A A A A A A A A A A A A A A A fb A A A
A Ac A A A A A A A A A A Afb A A A A A
A A A A A Afb A Y A A Afb A **sem** A
pred A

A	fb A Y fbA A A
A	AAfb A A A
A	A fb A A A A fb A A A

A

A A A A A A A A A A A A A A A A A A A
Ac Afb A Afb A A A A A A c A A A A

18. The Python Way of Working with a Listener

A Afb A fA A A A A A A A A A *mutatis mutandis* fA A
A A A A A A A A c A Afb A A A A Afb A A fBb A
A

A	A A
A	fb A A AA
A	fb A A A A
A	fb A A A A
A	fb A A A A
A	AA
A	fA A
A	AAAA AAG A
A	AAAA AA A
A	AAAA AA A
A	AAAA AA A
A	AAAA AA A
A	AA
A	AAAA AA A
A	AAAAA
A	AAAA AA A
A	AAAA AA A

A

A	fb A A A A
A	fb A A A A
A	fb A A A A
A	AA
A	A AA
A	AAA fa fbA A
A	AAAAAA fb A A A
A	AAAAAA fb A G c A
A	AA
A	AAA fa fbA AA
A	AAAAAA fb A
A	AA
A	AAA fa fbA AA
A	AAAAAA fb A
A	AAAAAA fb A A
A	AA
A	AAA fa fbA AA
A	AAAAAA A A A
A	AAAAAA A A A AA A A AAAAAA
A	AA
A	AAA fa fbA AAAAA
A	AAAAAA A A A
A	AAAAAA A A A
A	AA
A	AAA fa fbA AA
A	AAAAAA A A A
A	AA
A	AAAAAAfa A A A A A A AA
A	AAAAAAAA A A ☺ A
A	AAAA
A	AAAAAAfa A A A A A A AA

A	AAAAAAAAAA AA☹ A
A	AA
A	AAA fb fbA A
A	AAAAAA AA A fb A A Y A Y A

A A A fb A A fb A c A A A A A fb A A A
A A A A A A A A fb C A fb A A A A fb A A A A A A
A A A A A

19. Testing with Python

AW A A A A A A A fb A A A A A A A A A
A A A A A A A fb A A A fb c A

A	A A A A AA A A

A

A A A A A Ac Ac fb A A A A A A A A Ac fb A A
A A A A A A A A A A A A A fb A A A A A A A
A A Ac A A fb A A A A A A A A A A A A
A A A A A A A A A A A

A	A A
A	fb A A A A
A	fb A A A A
A	fb A A A A

A

A A Afb A A A A A A A A A A A

A AA Asymbol A A A A c A A A A A A

A	fb A A AA
A	fb A A A A
A	fb A A A A
A	fb A A A A
A	fb A A A A
A	A A
A	A A
A	AA
A	A A
A	AA
A	AAAA fbA fbA AAAAAAAAA
A	AAAAAAAAA AA AAAAAAAAA
A	AAAAAAAAA AA A
A	AAAAAAAAA AA A
A	AAAAAAAAAA
A	AAAAAAAAA fb AA A
A	AAAAAAAAA fb AA A
A	AA
A	AAAAAAAAA AAAAAAAAA
A	AAAAAAAAA AA fb A
A	AAAAAAAAA AAA
A	AA
A	AAAAAAAAA fb AA AAAAAAAAAAAAAAA
A	AAAAAAAAAA
A	AAAAAAAAA A A
A	AAAAAAAAAA
A	AAAA fbA fbA
A	AAAAAAAAA AA fb A A

A	AAAAAAAA A A AAAAAAAAAAAA A
A	AAAA
A	AAAAAAAA A A fb A
A	AAAAAAAA A A A
A	AAAAAAAA A AAAAAAAAAAAAAA
A	AA
A	AAAAAAAA A A A A A A A c A A AAAAA A
A	AAAAAAAA fb fb c A A
A	AA
A	AAAA fbA fbA
A	AAAAAAAA A A fb A
A	AAAAAAAA A A AAAAAAAAAAAA A
A	AAAA
A	AAAAAAAA A A fb A
A	AAAAAAAA A A A
A	AAAAAAAA A AAAAAAAAAAAAAA
A	AA
A	AAAAAAAA A A A A c A A A
A	AAAAAAAA fb fb c A A
A	AA
A	fbA A A A
A	AAAA A

A

A A A A A A A A A A A A A A A A A A
A Ac Afb A A A A fb A A A A A A A A
A A A A A A A A A A A Ac A A A

A

20. Parsing Markup

A A A A A Ac A A A A A A A A A A fb
c A AC A A A A c A A A A A A AYN A A
N AN A A A A fb Afb A A Afb A A A Ac A A A A A
A A A A A A Afb A Afb A fb
A A Ac A A A A A A A A A A
Aisland languages A A A A A A A A A A A A A A A A A A
fb A

G A A A A A Ac A A A A A A A A A A A A A
A A A Afb A A A A A A Afb A A A c A fbb A
A A A A A A A Ac A A A A A A
A A Ac A Afb A A A A

A A A A c Afb A A A A A A A A A A fb A A A A
A A A Ac A A A A A Ac A A A A fb A
fb Afb A A A A A A A A fb A A A A *lexical modes* A A
A A A A A A A A fb fb A A

21. Lexical Modes

A A A A A A A A A Ac A A A A A A

A	A AN A
A	AA
A	AAAAAAAAAAAAA A A A N CC AA
A	Y AAAAAAAAAAAA AA
A	AA
A	A A A A A
A	ACC A
A	AA
A	AAAAAAAAA A A A N AA
A	AAAAAAAAA A AA
A	R AAAAAAAAAAA AA
A	AAAAAAAAAAA A A A AA
A	AAAAAAAAAAAAA AA
A	AAAAAAAAAAAAAA A A AA
A	AA

A	fb A AAAAA AA

A

A A Afb A A A A A A fbb A A A fb A A A

A fbA A A c A A**You simply cannot define a lexical mode together with a parser grammar** A A A A A A A A A A A A A A c A

A A A A A A A A A A A A fb A A A fbACC A A A

A Ac A Ac A

A A A A A A A Afb Ac A A A A A A A c A A A A fb A A A A A A A A A A fbb A A A A A

A fb A A A A A fb AfbA A A A fb A A A A A A fb Ac A A A A Afb A A A*lexical modes* A A A A A A A A A A A A A A A A A A A Afb A A c A A

A A A A c A A A A fb A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A Ac A fb A A A A A fb A A A A R A A A A Ac A A A A A A A A A A**In the lexer grammar we need to define all tokens, because they cannot be defined later in the parser grammar.**A

22. Parser Grammars

A A A A A A fbA A A A A A A

A	A AN A
A	AA
A	AA W c N AA
A	AA
A	fb AAAAAAAA AA
A	AA
A	c AAAA A A AA
A	AA
A	AAA Y AA
A	AA
A	AA A A A AA
A	AA
A	AAAAA A A A c A A A A A A AA

A

A Afb A A A fb A A A A A A A A A fb A A A
A A A A A A A A A A A A A A Afb A A A A A
A A c A A A A A A fb A A A Afb A

A A A A A c A A A A

A A A A A A A A A A fb A A **content** A A A A A A A
A A A A A Afb A A A A A

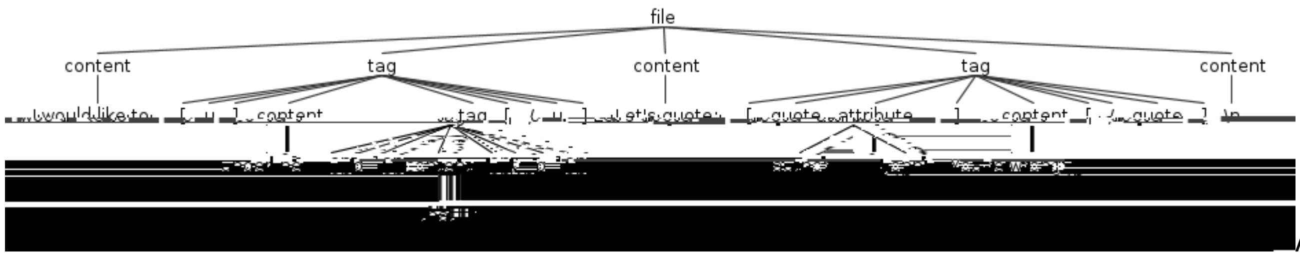
A A A A A A A A A A A A A A A A A A A A A A A
Ac A A A A A A A A A A

A A A fb A A A A A A Ac A A A A A
fbA A A A fb A A A A A AfbA A A A

A	AA A A c A A A A A A AA

C A A A A A A A A AfbA A A A A A fb A A A A
A A A A A A A A A A A fb A A c A A A
fb Afb A A A A A A A A A A fb A A A A A A AA
A A AA A A A A A fb A A A A A A A
A A A A A A A A A A A A

C fb A A A A A A A A A Afb AA A A



A

A A A A A A**element** A A A A fbA A A A A A Afb A
A A A A A Ac A A**tag** A A**content** A A A A A fb A A A A A A
A A A A A A fbA A A A fbA A A

A A Ac A fbA A A A A A A A A A A
A A A A A A fbA A A A A A A A fbA A A
A A A A fb A A A fbA A Ac A A A A A A Ac A A
A A A A A A A A A A A

Advanced

A A A A A A A fbA A A A A A A A A
A A A A A A A A A A A A A A A A fb A
A A A A A Ac A A A Ac A Afb A A A A A A

23. The Markup Project in Java

```
# use gradle to build the project
./gradlew compileJava
# if you are not using an IDE
# and you have defined the fatJar task as in the repository
./gradlew fatJar
java -jar .\build\libs\markup-example-gradle-all.jar
```

24. The Main App.java

25. Transforming Code with ANTLR

A	A
A	AAAA A
A	AAAA c A A G N G A A
A	AAAAA
A	AAAAA A
A	AAAAA A
A	AAAAA A
A	AAAAA A
A	AAAAA A A
A	AAAAA
A	AAAAA
A	AAAA A
A	AAAA c A A N A A
A	AAAAA
A	AAAAAAA Y A
A	AAAAAAAAA
A	AAAAAAA A A
A	AAAAA
A	A

A

Afb A AfbA A A A A A A A A A A A

A A A A A A Afb A Ac A A A A A A Afb A
A A A A A A A A A A A A A A A A A A
A A fb A A A A A A A A A A A A A A
A A A A A A A A A A Afb A A A A A A A
A A AfbA A A A A A A A A A A A A A A

26. Joy and Pain of Transforming Code

fb A A A A A A A A A A A A A A A
A Ac A A A

A	A
A	c A A N A AAAA
A	AAAAAAAAAA

A	A A A
A	A
A	A
A	A
A	c A A N A A
A	A
A	A fb A fb N G A
A	A A
A	A fb A A A
A	A A
A	A fb A A A
A	A A
A	A A
A	A
A	A A
A	A

A

C fb A A A A A A A A A A A G A A A A

Ac A A A A A A fb A A A A A

A A A A A A A A fb A Ac A Afb A A A A A A A A

A A tag A Ac A A A A Ac A A A A A A A A A

A A Ac A A Afb A A A A tag A content A A

A	A
A	c A A N A A
A	A
A	A A A A A
A	A A A A A A
A	A

A	Ac A
A	A A A A
A	c A
A	A A
A	A A A A
A	c A
A	A A
A	A c A c A
A	c A A c c c A
A	A A A A A A
A	A A A A A A A A
A	A c A A A
A	c A
A	A
A	
A	A A A
A	
A	fb AN A A A
A	
A	fb A A A
A	A A A
A	fb A A A
A	A A
A	
A	
A	A A A
A	
A	A
A	A

A

A A A A A A A Ac A A A A A A

A A A A A A Ac A A A A A A Ac A

A A A A A A fbA AD A A A fbA A A A A A A A A
A A A Ac A A A A A A A A A A A A A
fb A

Afb A A A A A A A A A fb A A A fb A A A
A A ffb A A A A A fb A Ac A A A ACC A A A
c A A A A fb A Afb A N AN A A A A A A c A
fbA N A A

A	A A

A

A A A A A A A A A A A A A A A Ac A A A A
c A A A A A A A A A Afb Ac A A A A A A A A
A A A

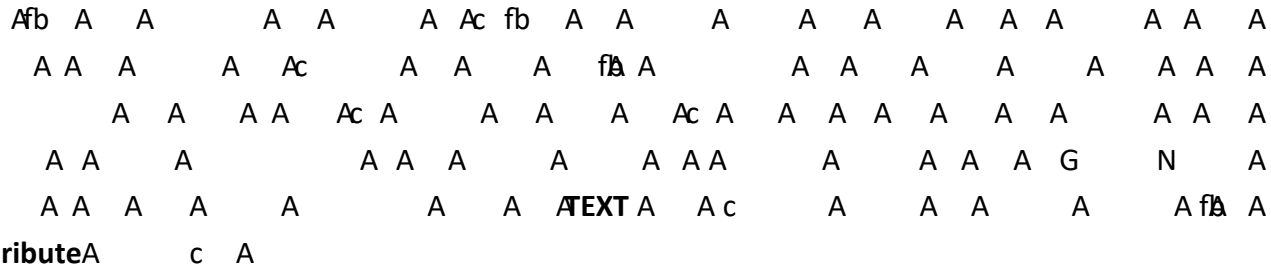
```java
    // private variables inside the class AppTest
    private MarkupErrorListener errorListener;
    private MarkupLexer markupLexer;

    @Test
    public void testText()
    {
        MarkupParser parser = setup("anything in here");

        MarkupParser.ContentContext context = parser.content();

        assertEquals("",this.errorListener.getSymbol());
    }

    @Test
    public void testInvalidText()
    {
        MarkupParser parser = setup("[anything in here");

        MarkupParser.ContentContext context = parser.content();

        // note that this.errorListener.symbol could be empty
        // when ANTLR doesn't recognize the token or there is no error.
        // In such cases check the output of errorListener
```

A	assertEquals("[",this.errorListener.getSymbol());A
A	}A
A	A
A	@TestA
A	public void testWrongMode()A

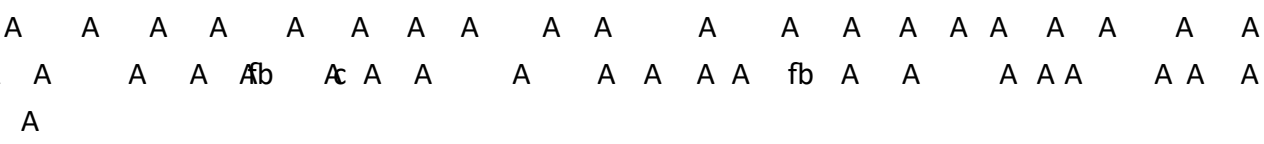
A

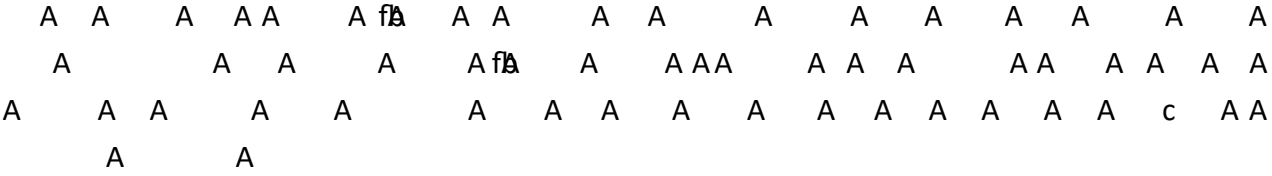
```
        A
        assertEquals("",this.errorListener.getSymbol());A
    }A
A
    @TestA
    public void testInvalidAttribute()A
    {A
        MarkupParser parser = setup("author=/\"john\"");A
        // we have to manually push the correct modeA
        this.markupLexer.pushMode(MarkupLexer.BBCODE);A
        A
        MarkupParser.AttributeContext context = parser.attribute();
A
        A
        assertEquals("/",this.errorListener.getSymbol());A
    }A
```

A
A

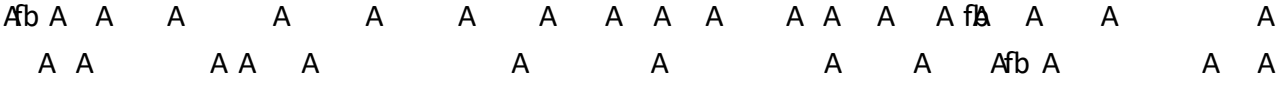


attributeA c A





28. Dealing with Expressions



left-recursive rules.

ANTLR4 can create a similar structure automatically, so we can use a much more natural syntax

A	AA A A A
A	A A A A
A	A N C A
A	A

A

A A A A A A A A A A fb A A A A A A

AC A A A A A A A A A A A A A A A A

A A A A

29. Parsing Spreadsheets

A A A A A A A A A A A A A Ac AA A A fb A

A A A A A A A A A A A A A A A A fb A

A

A	A A
A	AA
A	A A A A A
A	A A A A A
A	A A N A A c A
A	A AA A A AA A
A	A N A A A A fb A
A	A N C A A
A	A A A
A	A
A	AA
A	fb A AA AA
A	fb A AA A AA
A	AA
A	A AA
A	A A AA
A	A A AA
A	A A AA
A	N A A AA
A	AA

A	AAAAAAAAAAAAAAAA A AA
A	AA
A	N AAAAAAAAAAAAAA AA
A	AA
A	N C AAAAAAAAAAAA A A AA
A	AA
A	AAAAAAAAAAAA A A

A

A A A A A A A AAb A A Ac A A Aab A c A

A A

A A A A A

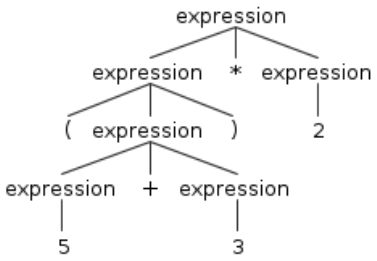
A A aba A A A

A A A A A

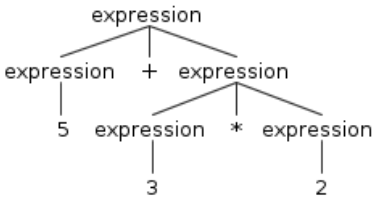
A A Aab Ac A A A A A A A A A A A A A A

A aba A aba A A A A A A A A aba A A A

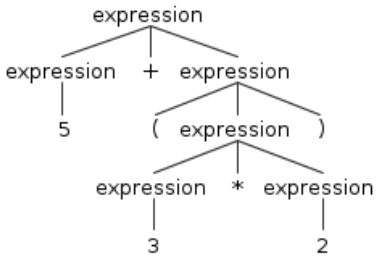
A A A



(5+3)*2



5+3*2



5+(3*2)

A A A A A *Alabels* A A A A A A A A fb Aab Aab A A

A A A A A Ac A A A A A

A c A A A A A aba A A Aab A A**expression**A A

A A A A A A A abb Ac A A A A c A A A A

A A A A A A A A A A aba A A A Aab A A A A

A A A A A fbAab A A A A A A A A A A A A A A A

A A*associativity* A A A A A A A A A A A*left-associativity,* A A A A

fb A A A A A*right-associative* A A A A A A A A A A

Another way to look at this is: if there are two expressions of the same type, which one has the precedence: the left one or the right one?A A A A A A A A A

A	AAAAAAAAAA A AA A A
A	AAAAAAAAAA A AA A A
A	AA
A	AAAAAAAAAA A AA A
A	AAAAAAAAAA W A AA A W A
A	AAAAAAAAAAA
A	AAAAAAAAAA W A
A	AAAAAAAA A
A	AAAA A
A	A

A

A A A A A Afb A A fbA A A A A A A A A A A A A A A A fbA A A A A A AG A A A A A A A A A A A A A A

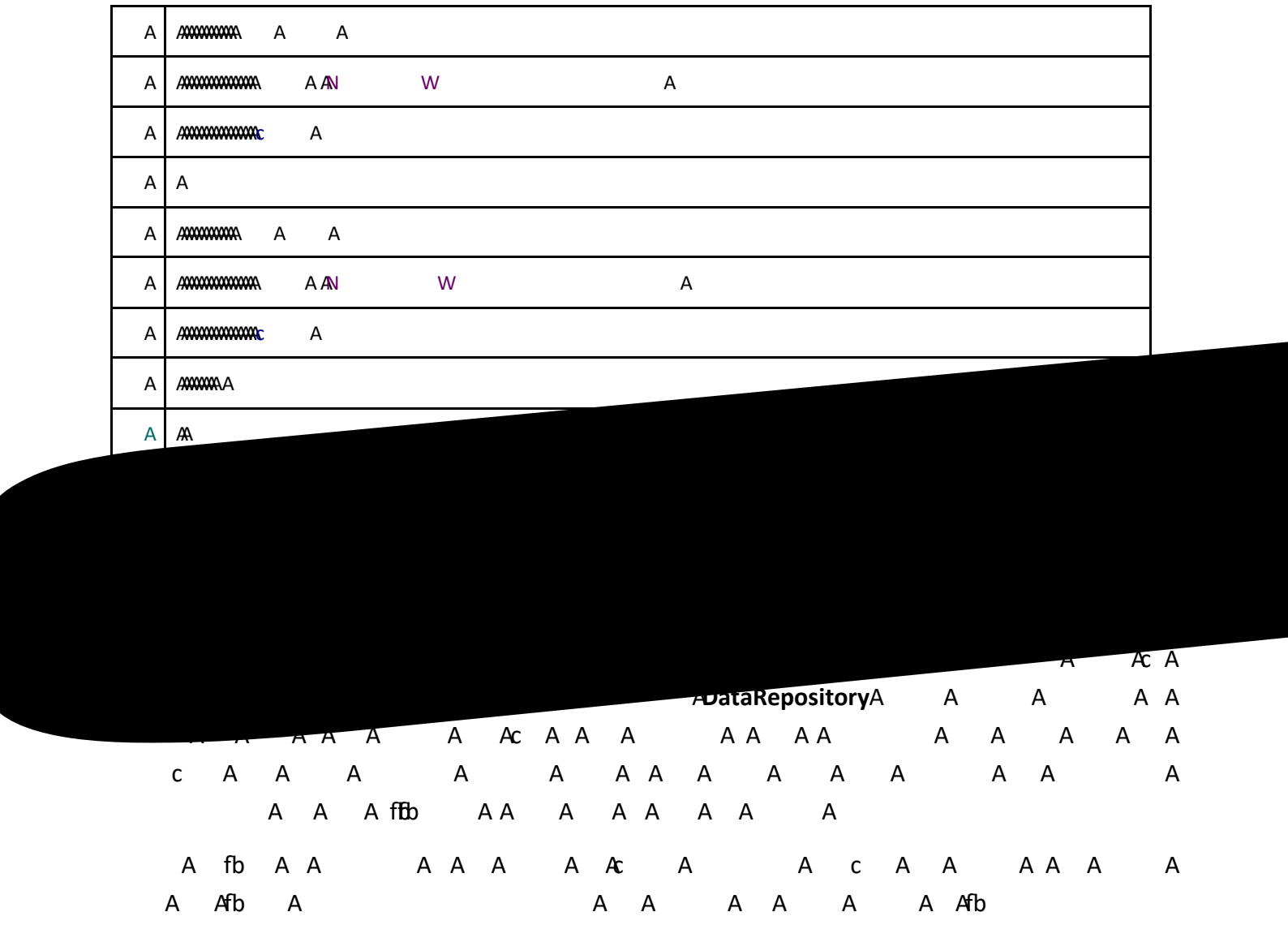
A A A A A A A A A A A A A A A A A fbA A A A fbA A A A A AAfb A

31. Excel is Doomed

A A A A AA A A A Afb A A*Spreadsheet*A A

A	c A A W AA C W c A
A	A
A	AAAA A A A AA A A
A	AA
A	AAAA c A A c AW A A
A	AAAA AAAAAAAAAAAA
A	AAAAAAA A c N C A c fb A
A	AAAA A
A	AA
A	AAAA c A A c AW A A
A	AAAA A
A	AAAAAAA A A A A
A	AA

A	A A A
A	A A
A	A
A	A c A A c A A A
A	A A
A	A AV A
A	A A
A	A
A	A c A A c A N N A A
A	A A
A	A c A fb A A
A	A c A A A
A	A c A A A A
A	A
A	A fb A A A
A	A A fb A A A
A	A A fb A A A
A	A
A	A A
A	A A
A	A
A	A
A	A c A A c A G G A A
A	A A
A	A A A N A
A	A c A A A A
A	A
A	A
A	A A



```
A   public void testExpressionPow()A
A   {A
A       setup("5^3^2");A
A   A
A       PowerExpContext context = parser.expression() as PowerExpContext;A
A   A
A       CommonTokenStream ts = (CommonTokenStream)parser.InputStream;    A
A   A
A       Assert.Equal(SpreadsheetLexer.NUMBER, ts.Get(0).Type);A
A       Assert.Equal(SpreadsheetLexer.T__2, ts.Get(1).Type);A
A       Assert.Equal(SpreadsheetLexer.NUMBER, ts.Get(2).Type);A
A       Assert.Equal(SpreadsheetLexer.T__2, ts.Get(3).Type);A
A       Assert.Equal(SpreadsheetLexer.NUMBER, ts.Get(4).Type); A
A   }A
A   A
A   [Fact]A
A   public void testVisitPowerExp()A
A   {A
A   
```


A

```
A A
A [Fact]A
A public void testWrongVisitFunctionExp()A
A {A
A     setup("logga(100)");A
A A
A         FunctionExpContext context = parser.expression() as FunctionExpContext;A
A     A
A     SpreadsheetVisitor visitor = new SpreadsheetVisitor();A
A     double result = visitor.VisitFunctionExp(context);A
A A
A     CommonTokenStream ts = (CommonTokenStream)parser.InputStream;A
A A
A     Assert.Equal(SpreadsheetLexer.NAME, ts.Get(0).Type);A
A     Assert.Equal(null, errorListener.Symbol);A
A     Assert.Equal(0, result);A
A }A
A A
A [Fact]A
A public void testCompleteExp()A
A {A
A     setup("log(5+6*7/8)");A
A A
A     ExpressionContext context = parser.expression();A
A A
A     SpreadsheetVisitor visitor = new SpreadsheetVisitor();A
A     double result = visitor.Visit(context);A
A A
```

```
A       Assert.Equal("1.0107238653917732",
    result.ToString(System.Globalization.CultureInfo.GetCultureInfo("en-
    US").NumberFormat));
```

Catchall Rule

Afb A A A AANY A A A A A A A A A Afb Afb A

A	AAAA

A

A A A A A A Ac A A A A A fbA A A A A A A A
Afb A A A A A A A A A A A A A A A A
A A A A A A A A A A A A A A A A
fb A A A A A A A A A A A AfbA A A A A A A A
A A A A A

Channels

A A A A A A A A A c A*channels* A A A A A A A
A A A A A A A A Afb A A A A fbA A A A
A A A A A A A A A A AC A A A A A A A
A A A A Afb A AfbA A A A A A A A A
A A A A A*channels* A A A A A A A A A A Ac A
A A A A fbA A A A A fbA A A A

A	AA R N AA
A	A A A A A A A
A	N N AA A A A R N AA

Rule Element Labels

A A A A fbA c A A A A A A ffib A A fbA A A A
A Ac A A A A A fb A A Ac A A A fbA A A A A A
A A A fbA A A Afb A A c A Ac A A A A A

A	AA fb A A AA

A

A A**left** A A**right** A Ac Afb A A A A A A A fbA
A A A fb A A A A A A A A

Problematic Tokens

A A A A A c A A A A ffib A A A fbA A A A A
c A A A c A A A A Ac A Ac Afb Ac fbA
A A A A A A

A	Ac fbA A A A AA A A fbaA fb Ac A Ac A
A	AAA A A
A	A A A A fb AA AA A fba A
A	A A A

A

A A A fba fb A Ac fbA A A A AA A c A Ac A A
C A A A A A fb A A A A fb A A Ac fbA Afb A
A A A A A A A A AA A A fba A A c A A
c A A A A A A c A fba A A A A A A A
A A A A fb A Ac fbA A A A A A A Ac A
A A A A A A A A A fba A A Ac A Afb A A A

A	Afb A A
A	G AA A
A	AA A G A N C A
A	A A A
A	AA A G A G A N C A

A

34. Conclusions

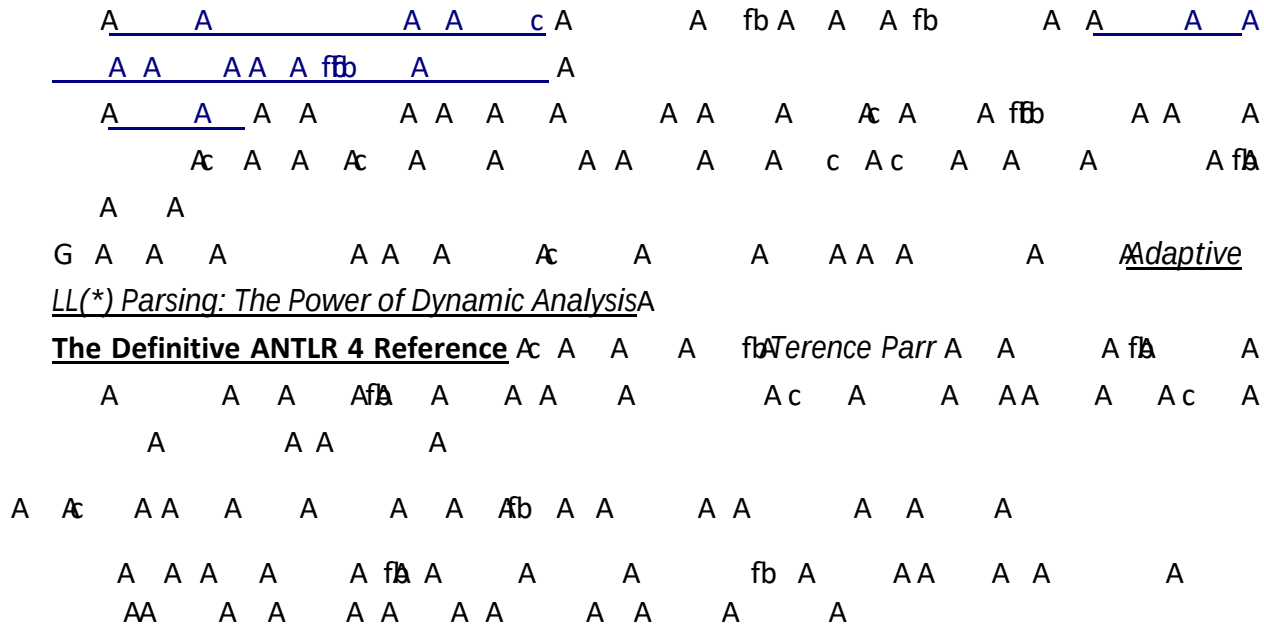
A A AA A A

- A AA A AA A
- A A A A A A A
- A A A A A A A A A A A A A A A
- Afb A A fba c A A A A A A A A A A A
- A A A A
- A A A A A

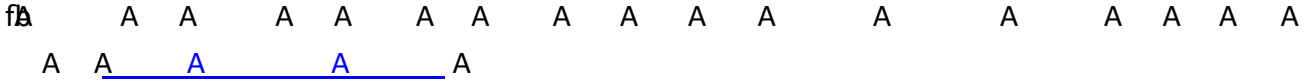
A A A A A A A A A A A A AA A A A A A A A
A Ac A A A A Ac A A A A A A

A A Afba A A A fb Ac A A

A A A c A A A A A A A A A
A fbb A A c A A A A A A A A A A A A fba A
A A A A A A A A A A A



Why program by hand in five days what you can spend twenty-five years of your life automating?



We would like to thank Bernard Kaiflin for having revised the document and helped us improving it.

We would like to thank: Brasilio Castilho, Andy Nicholas for having spotted errors and typos in the article.

We worked quite hard to build the largest tutorial on ANTLR: the mega-tutorial! A post over 14.000 words long, or more than 70 pages, to try answering all your questions about ANTLR. Missing something? Contact us and let us now, we are here to help.