SEARCH-BASED AMBIGUITY DETECTION IN CONTEXT-FREE GRAMMARS

NAVENEETHA VASUDEVAN, LAURENCE TRATT
King’s College London, Strand, London, WC2R 2LS, United Kingdom

AMBIGUITY PROBLEM
1. Ambiguity is an undecidable problem [1]
2. Multiple parse trees are undesirable

\[ E \rightarrow E + E \mid E \ast E \]

\[ 4 + 3 \ast 2 \]

Parse Tree 1  Parse Tree 2

\[ \begin{align*}
E & \quad + \\
E & \quad 2 \\
4 & \quad 3 \\
\end{align*} \]

\[ \begin{align*}
E & \quad E \\
E & \quad + \\
E & \quad 4 \\
3 & \quad 2 \\
\end{align*} \]

Value = 14 Value = 10

CFGs

Real-world language grammars

Random grammars

Context-free grammars

SinBAD

Grammar artifacts

Sentence Generator

Earley parser

Parsed output

Ambiguity?

No

Time exceeded?

Yes

Results

Random grammars

Altered real-world grammars

EXPERIMENT
To understand how well our search-based approach uncovers ambiguity. We evaluate SinBAD against two tools:

1. ACLA - uses an approximation technique
2. AmbiDexter - uses a hybrid approach (grammar filtering + search)

Grammar sets include:

1. Random grammars (tested for 10, 30, 60, & 90s)
2. Altered real-world grammars [2] (tested for 180 & 300s)

FUTURE DIRECTIONS
• To add more tools to the study and perform a larger experiment with more real-world-esque grammars
• To develop a CFG equivalent of the work on random generation of automata [3]

SOURCE CODE
The source code for SinBAD is available at http://github.com/nvasudevan/sinbad

REFERENCES