

SIGMA: Systematic Island Grammar forMation Approach

Merging Grammars

Introduction

Motivation—
Research Goal—
Research Question—

Approach

- Steps
1. Parse Grammars
 2. Trivially Merge Grammars
 3. Normalize Grammar
 4. Measure Production Similarities
 5. Merge Most Similar Productions
 6. Repeat Steps 3–5 Until Max Similarity is Below a Threshold
 7. Output Grammars

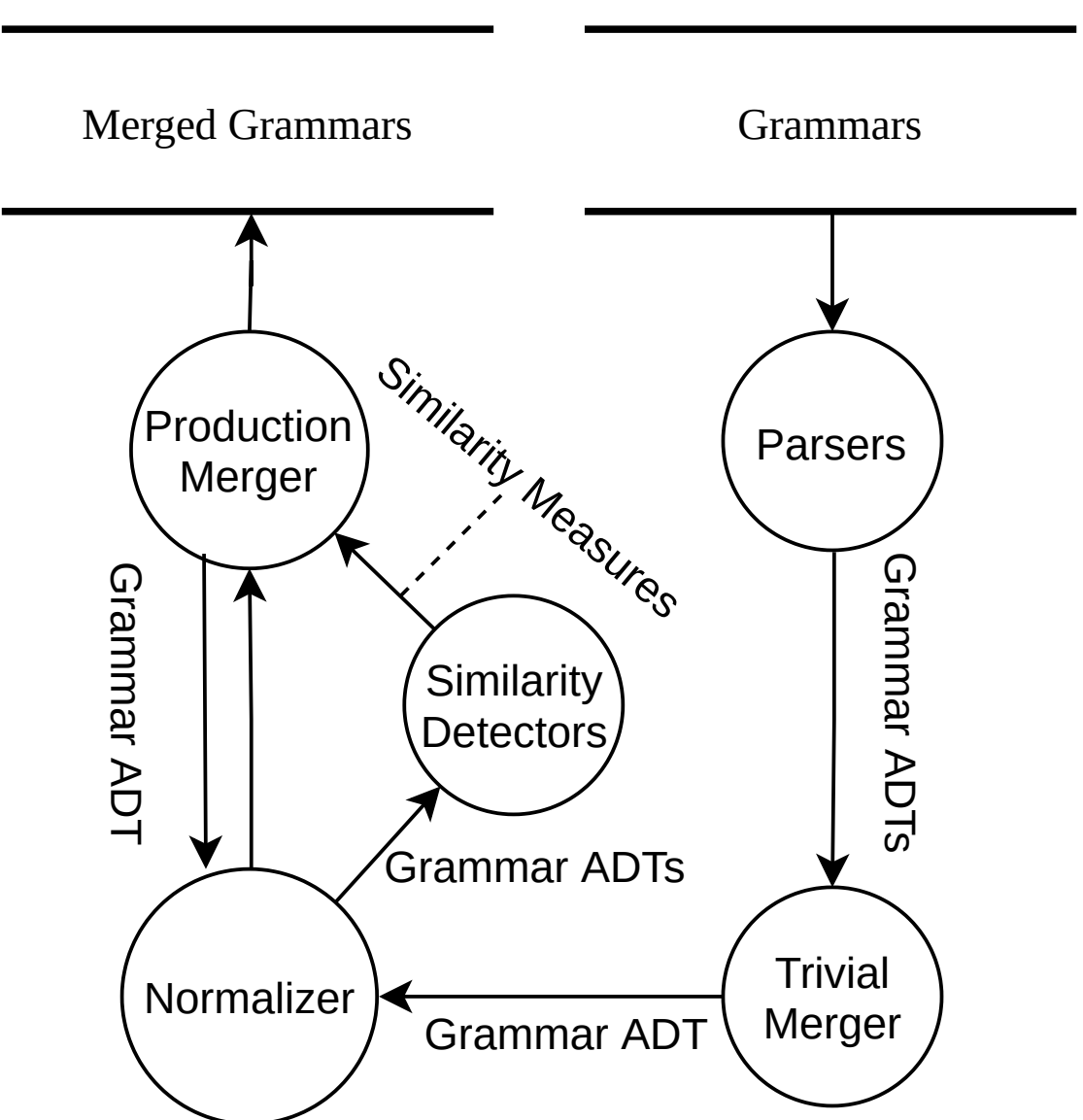


Fig. 1: Merge Process

- Data Model
- Object Based
 - Right Hand Side of Productions is an Object
 - Constructed via Transformation of Grammar's Abstract Syntax Tree
 - Converted to Text via Visitor

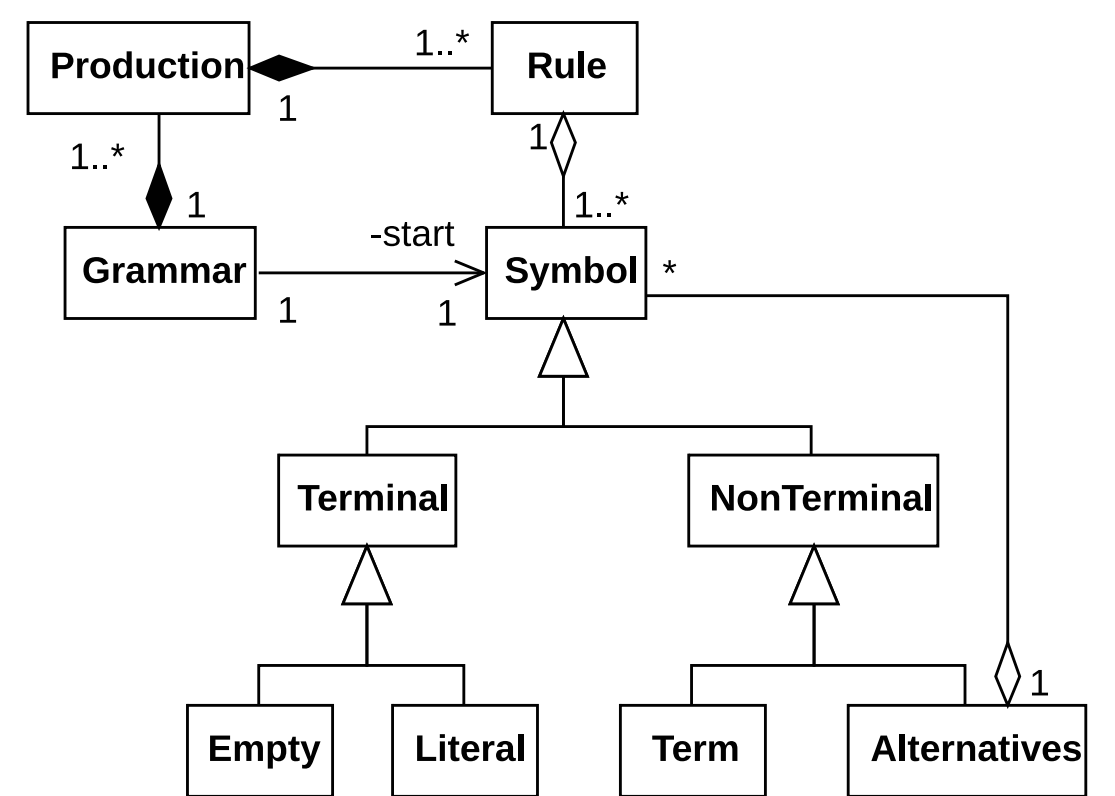


Fig. 2: Data Model

Measuring Production Similarity

Productions P_a, P_b like $A 'a' B$

Productions P_a, P_b like $A | 'a' | B$

$$\frac{2|LCS(P_a, P_b)|}{|P_a| + |P_b|}$$

$$\frac{2|P_a \cup P_b|}{|P_a| + |P_b|}$$

LCS returns the longest common subsequence.

Normalization

Normalizes grammars so that all rules match one of two forms:

$P_1 \rightarrow A 'a' B$

$P_2 \rightarrow A | 'a' | B$

Experimental Design

Results

Discussion

Conclusions

References

Acknowledgements

This research is supported by funding from the Ronald E. McNair Post Baccalaureate Achievement Program at Idaho State University, which is sponsored by the Department of Education (P217A170169).