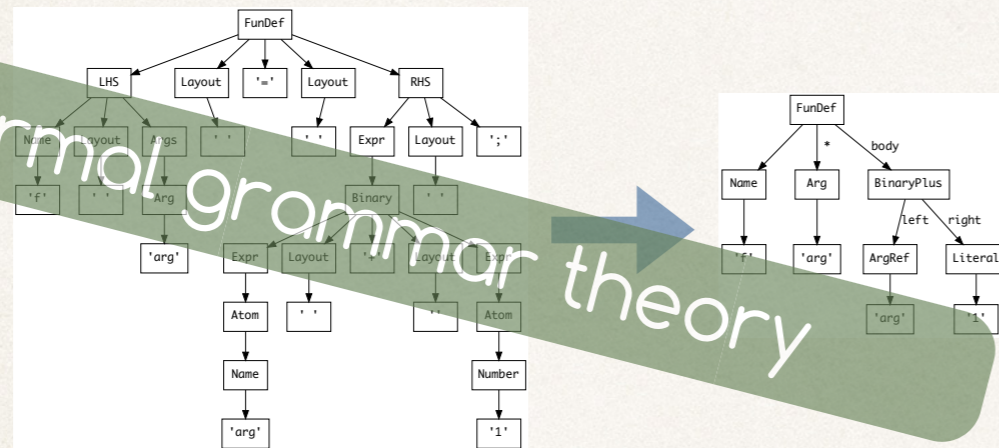


Formal Foundations for Semi-parsing

CSMR-WCRE 2014: Early Research Achievements

Vadim Zaytsev, Universiteit van Amsterdam, The Netherlands

Parsing:



(hard to automate for legacy code, unknown dialects, language cocktails, ...)



Lexical analysis:

- * grep
- * sed
- * awk
- * perl
- * regular expressions
- * search & replace
- * ...

string manipulation

(easy to set up, very error prone)

Boolean grammars:

$A ::= B$

$A ::= B C D$

$A ::= B | C D | E F$

$A ::= (B | (C \& D) | E F) \& \neg X$

(conjunction, negation, disjunction, seq. composition of symbols)

Parsing schemata:

- * deduction system
 - * item set (parse forest specs)
 - * hypotheses set
 - * deduction steps
- * inference relation \vdash
- * ...

Comfortaa by Johan Aakerlund (SIL Open Font License)
Heuristica by Andrej Panov (SIL Open Font License)
Chinese black dragon by ANGELUS (CC-BY-SA; GNU FDL)
Boolean grammars by Alexander Okhotin (CSR 9, 2013)
Parsing schemata by Klaas Sikkel (Springer, 1997)