Guided Convergence

Vadim Zaytsev, CWI, Amsterdam
http://grammarware.net

language comparison

Language X

e.g. Java as defined in the "Java Language Specification" by Gosling, Joy, Steele, Bracha

(intended the same)

?

Language Y

e.g. Java as actually accepted by the javac compiler

defines

defines

grammar identity

Grammar X

the grammar in a broad sense that defines the language (e.g., in EBNF, XSD, UML, Ecore)

=

Grammar Y_1

this grammar in a broad sense identical to the grammar X to the smallest detail

nominal equivalence

bijectional mapping of production rules

Grammar Y_2

this grammar uses the same nonterminals as grammar X and just as many production rules that we can map to rules of X

F ::= N "(" {A ","}* ")" ("=" B)?;

vs.

F ::= N A* B?;

structural equivalence

name mapping for nonterminals

Grammar Y_3

nonterminals of Y do not match with nonterminals of X, but we think they should

F ::= N "(" {A ","}* ")" ("=" B)?;

vs.

G ::= Id D? P+ W;

extracting basic structure from each production rule

extracting basic structure from each production rule

≈

X Signatures

I,*,?

Y_3 Signatures

I,?,+,I

case study:

ANTLR, TXL, ADT, DCG, Rascal, SDF, EMF, JAXB, Java, Python, XML Schema

converged