



What Have We Done About the  
Unnecessary Diversity of  
Notation for Syntactic Definitions



Diversity of  
notation

# Diversity of syntactic notation

- Software languages and grammarware are everywhere
- Grammars are extracted from documentation
- Every language document uses its own notation

## ANSI Smalltalk

`<<class methods>> ::= <method definition>*`

`<<class initializer>> ::= <initializer definition>`

# Diversity of syntactic notation

ISO/ECMA C#

*member-access:*

*primary-expression . identifier type-argument-list<sub>opt</sub>*

*predefined-type . identifier type-argument-list<sub>opt</sub>*

*qualified-alias-member . identifier type-argument-list<sub>opt</sub>*

*predefined-type: one of*

bool byte char decimal double float int long object sbyte short string uint ulong ushort

AL/Algol-58

$\langle de \rangle ::= \langle id \rangle \overline{\text{or}} \langle integer \rangle \overline{\text{or}} \langle id \rangle [\langle exp \rangle]$

$\langle del \rangle ::= \langle de \rangle \overline{\text{or}} \langle del \rangle , \langle de \rangle$

$\langle \text{switch declar} \rangle ::= \text{switch } \langle id \rangle := (\langle del \rangle)$

# Diversity of syntactic notation

## Google Dart

**normalFormalParameters:**

normalFormalParameter (‘, ’ normalFormalParameter)\*

;

**namedFormalParameters:**

‘[’ defaultFormalParameter (‘, ’ defaultFormalParameter)\* ‘]’

;

## IEEE Scheme

(cond <clause1> <clause2> ...) syntax

*Syntax:* Each <clause> shall be of the form

( <test> <expression> ...)

# Diversity of syntactic notation

Ada 83

subtype\_declaration ::= **subtype** identifier is subtype\_indication;

subtype\_indication ::= type\_mark [constraint]

type\_mark ::= *type\_name* | *subtype\_name*

ISO/IEC Eiffel

Note\_name  $\triangleq$  Identifier ":"

Note\_values  $\triangleq$  {Note\_item ", " ...}+

Note\_item  $\triangleq$  Identifier | Manifest\_constant

# Diversity of syntactic notation

Scala v.2.7

Block ::= {BlockStat semi} [ResultExpr]

BlockStat ::= Import

| ['implicit' | 'lazy'] Def

| {LocalModifier} TmplDef

| Expr1

|

W3C XPath2

[50] SequenceType ::= ("empty-sequence" "(" ")")  
| (ItemType OccurrenceIndicator?)

[52] ItemType ::= KindTest | ("item" "(" ")") |  
AtomicType



# Notation Specification



# Confix metasymbols

- start grammar / end grammar
- start nonterminal / end nonterminal
- start group / end group
- start star repetition / end star repetition
- start star separator list / end star separator list
- ...

# Infix metasymbols

- terminator metasymbol
- defining metasymbol
- definition separator metasymbol
- concatenation metasymbol
- exception metasymbol
- ...

# Postfix & prefix metasymbols

- postfix optionality metasymbol
- postfix star repetition metasymbol
- postfix star repetition metasymbol
- start one line comment metasymbol
- ...

# Conventions

- whitespace reliability
- indentation
- glue consecutive terminals
- uppercase / lowercase / camelcase / mixed case
- ...

# Predefined sets

- masked terminals (“” as “x” with x being ”)
- nonterminals may contain (space? underscore? slash?)
- built-in nonterminals (string? identifier? number?)
- ...



Conclusion

# Results

- Analysed 69 sources with 42 different notations
- Formulated a notation specification
- Prototyped as EDD = EBNF Dialect Definition
- Created Grammar Hunter, an EDD-based recovery tool
- Extracted 64 grammars

# Opportunities

- Grammar extraction from language documentation
  - Zaytsev, Notation-Parametric Grammar Recovery, LDTA 2012
- Generating correct documentation
  - Zaytsev, Lämmel, A Unified Format for Language Documents, SLE 2010
- Syntactic notation evolution
  - Zaytsev, Language Evolution, Metasyntactically, BX 2012
- Linking documentation to compilers
- Grammar readability
- Documentation completion
- Grammar comparison



# Credits

- **URBAN JUNGLE** — font by KC Fonts  
(free for non-commercial use)
- **A drippin' marker** — font by Wickhop (freeware)
- **Wickhop Handwriting** — font by Wickhop (freeware)
- "What Can We Do About the Unnecessary Diversity of Notation for Syntactic Definitions?" — a paper by Niklaus Wirth
- Photographs — all made by Vadim Zaytsev

VADIM ZAYTSEV  
GRAMMARWARE.NET  
SWAT, CWI

Questions?

**DNE  
WAS  
HERE**