

13. Context-Sensitive Grammars

This series of tutorials is based upon work from COST Action
Multi3Generation CA18231, supported by COST
(European Cooperation in Science and Technology).

COST (European Cooperation in Science and Technology) is a funding agency for research and innovation networks. Our Actions help connect research initiatives across Europe and enable scientists to grow their ideas by sharing them with their peers. This boosts their research, career and innovation, cf. www.cost.eu

Context-Sensitive Grammars

- Context-sensitive grammars are equivalent to Type 1 Generative Grammars in Chomsky-Schützenberger's hierarchy
- Context-sensitive grammars can be more powerful than regular and context-free grammars.
- They can recognize any context-sensitive language, are very well-adapted to describe distributional and agreement constraints between linguistic units, and can be processed by linear bounded automata.
- Lexical Functional Grammars (LFG) are equivalent to context-sensitive grammars.

Context-Sensitive Grammars

- In NooJ, context-sensitive grammars are context-free grammars + variables and constraints.
- Context-sensitive grammars have two components: the description of sequences and the description of agreement constraints.

Type 1 : Context-Sensitive Grammars

A grammar that recognizes words that contains the same number of “a”, “b” and “c”

NooJ

File Edit Lab Project Windows Info GRAMMAR

aⁿ bⁿ cⁿ.nog

French /French syntactic grammar.

Diagram illustrating the grammar structure for aⁿ bⁿ cⁿ:

The diagram shows three components labeled A, B, and C, each enclosed in red parentheses and connected by minus signs. Component A contains a box labeled 'a' with a self-loop. Component B contains a box labeled 'b' with a self-loop. Component C contains a box labeled 'c' with a self-loop. These components are followed by a final state symbol (a circle with a cross) and two constraints: <\$B\$LENGTH=\$A\$LENGTH> and <\$C\$LENGTH=\$A\$LENGTH>.

Contract for: aⁿ bⁿ cⁿ

Check

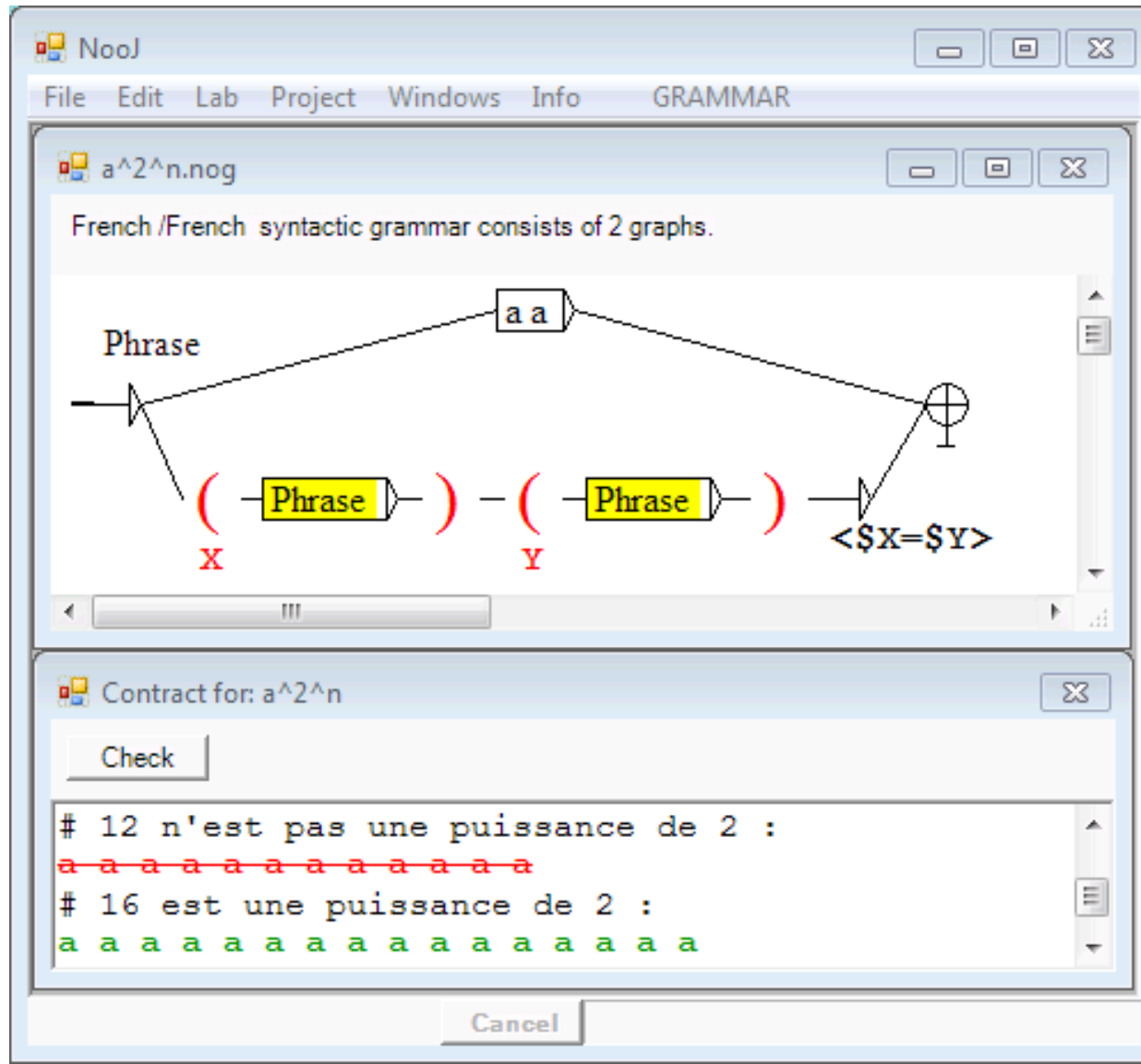
NooJ
Enter examples, *counter-examples and #comments

a a a b b b c c c
~~a a a b b c c c~~

Cancel

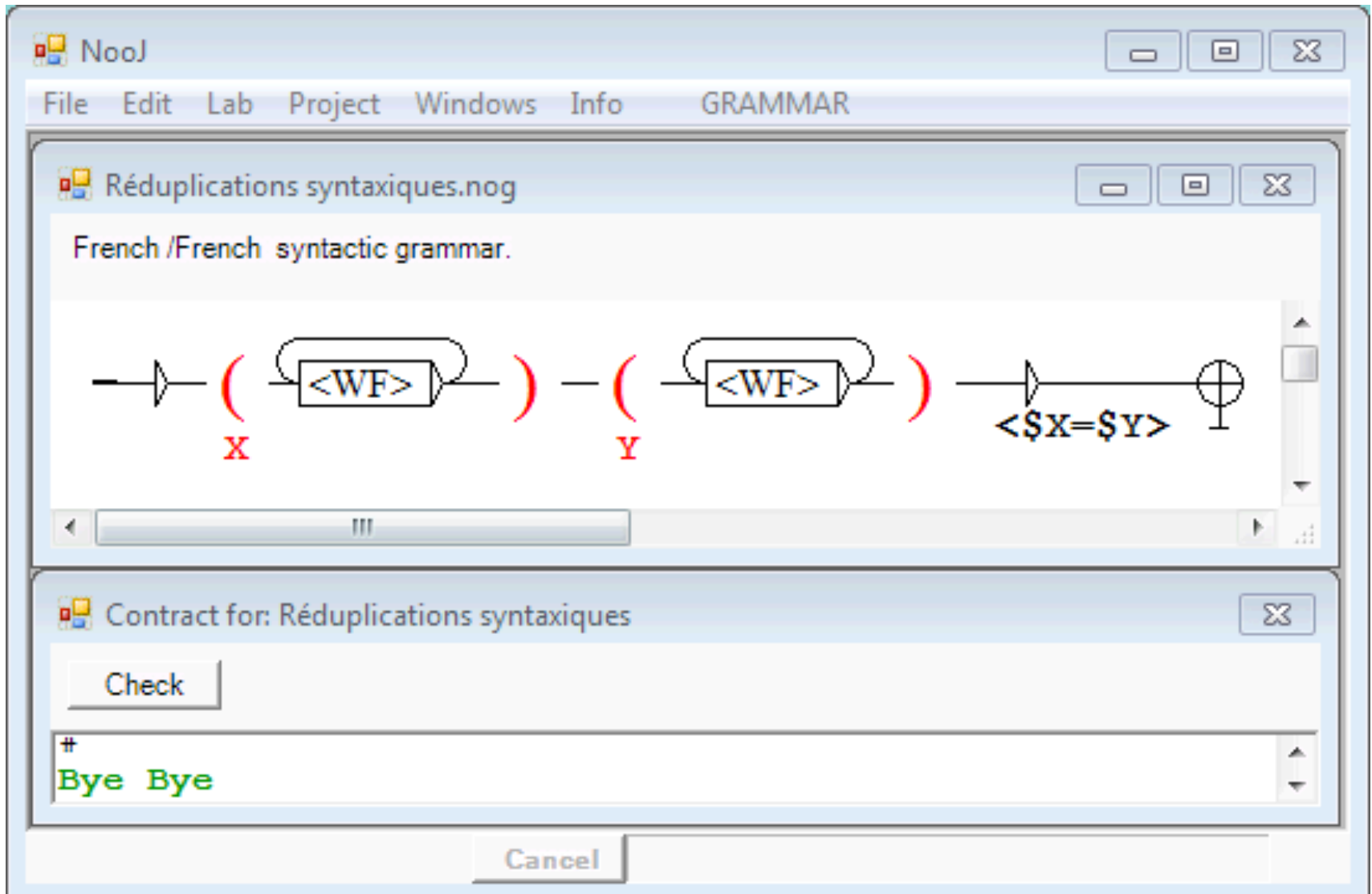
Type 1 : Context-Sensitive Grammars

A grammar that recognizes words constituted by a number of "a" that is a power of 2, e.g., 2, 4, 8, 16...



Type 1 : Context-Sensitive Grammars

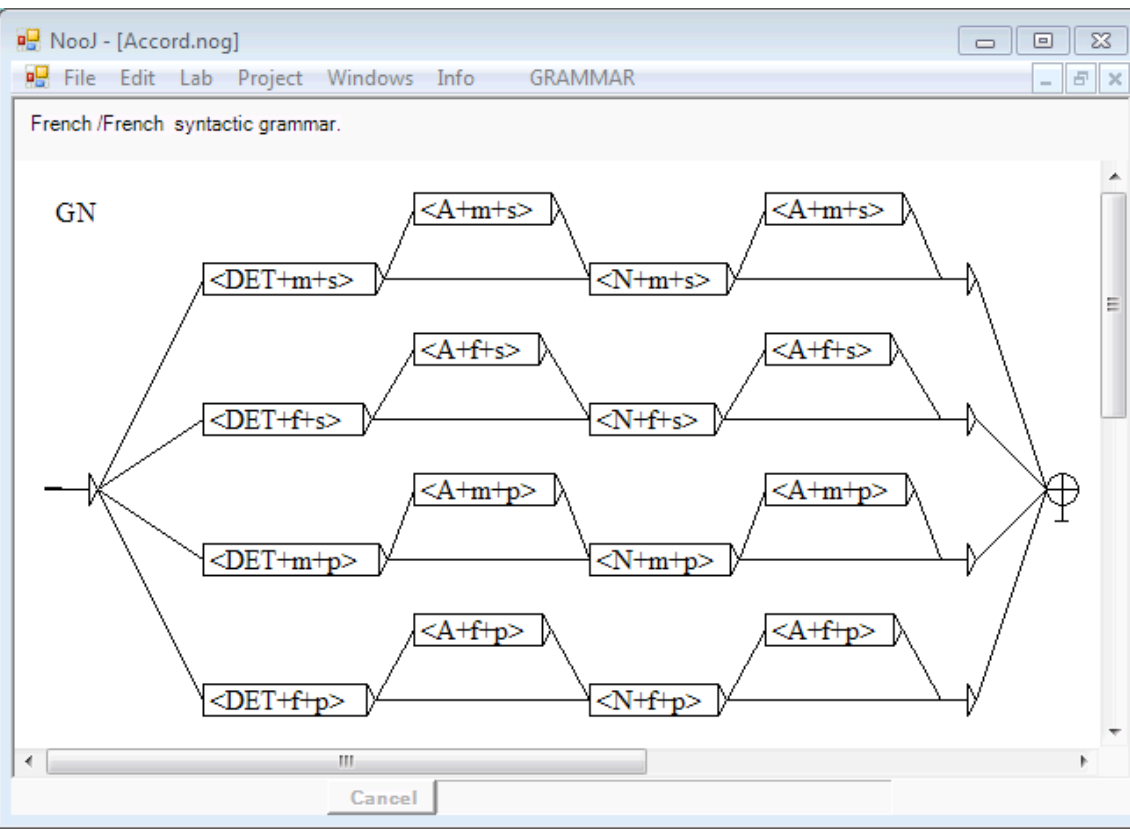
A grammar that recognizes reduplications, e.g., bye bye, chouchou, papa



Context-Sensitive Grammars

- Example of a grammar that describes the agreement in number and in gender in some French noun phrases:

$\langle \text{DET}+\text{m}+\text{s} \rangle (\langle \text{A}+\text{m}+\text{s} \rangle | \langle \text{E} \rangle) \langle \text{N}+\text{m}+\text{s} \rangle (\langle \text{A}+\text{m}+\text{s} \rangle | \langle \text{E} \rangle) |$
 $\langle \text{DET}+\text{f}+\text{s} \rangle (\langle \text{A}+\text{f}+\text{s} \rangle | \langle \text{E} \rangle) \langle \text{N}+\text{f}+\text{s} \rangle (\langle \text{A}+\text{f}+\text{s} \rangle | \langle \text{E} \rangle) |$
 $\langle \text{DET}+\text{m}+\text{p} \rangle (\langle \text{A}+\text{m}+\text{p} \rangle | \langle \text{E} \rangle) \langle \text{N}+\text{m}+\text{p} \rangle (\langle \text{A}+\text{m}+\text{p} \rangle | \langle \text{E} \rangle) |$
 $\langle \text{DET}+\text{f}+\text{p} \rangle (\langle \text{A}+\text{f}+\text{p} \rangle | \langle \text{E} \rangle) \langle \text{N}+\text{f}+\text{p} \rangle (\langle \text{A}+\text{f}+\text{p} \rangle | \langle \text{E} \rangle)$



Four quasi-identical terms for French...

But 30 quasi-identical terms for some slavic languages (2 numbers, 3 genders, 5 cases).

Context-Sensitive Grammars

- Example of a Context-Free Grammar that only describes the structure of some French noun phrases:

$$\text{GN} = \langle \text{DET} \rangle (\langle \text{A} \rangle | \langle \text{E} \rangle) \langle \text{N} \rangle (\langle \text{A} \rangle | \langle \text{E} \rangle)$$

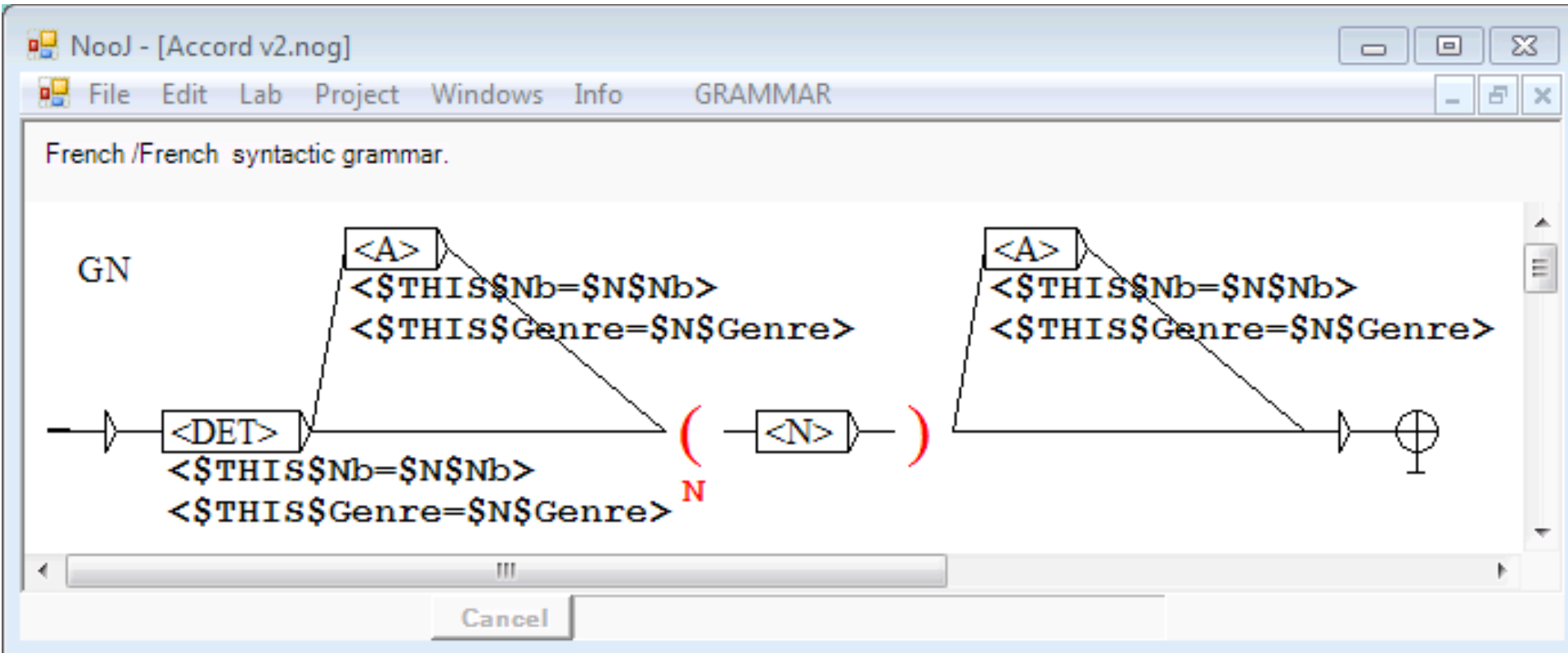
- We then add agreement constraints for gender and number:

$\langle \$\text{THIS}\$ \text{Gender} = \$\text{N}\$ \text{Gender} \rangle$

$\langle \$\text{THIS}\$ \text{Number} = \$\text{N}\$ \text{Number} \rangle$

- Variable $\$ \text{THIS}$ refers to the current ALU; variable $\$ \text{N}$ refers to the noun.

Context-Sensitive Grammars

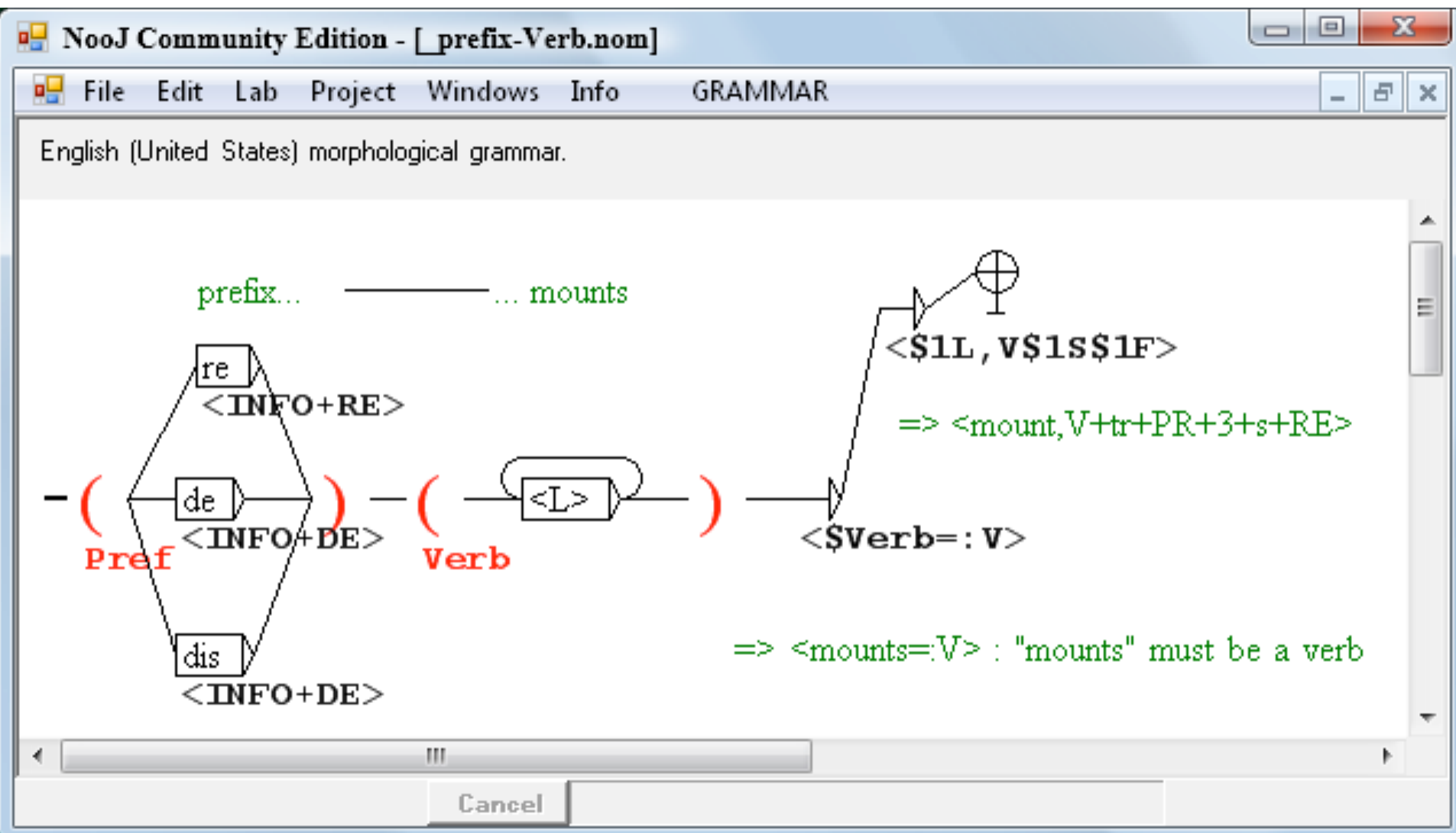


- The noun **<N>** is stored in variable **\$N**
- Special variable **\$THIS** refers to the current ALU
- Variable **\$N\$Nb** refers to the “NB” property of the noun; **\$THIS\$Nb** refers to the “NB” property of the current ALU;
- Variable **\$N\$Genre** refers to the “Genre” property of the noun; **\$THIS\$Genre** refers to the “Genre” property of the current ALU.

A morphological context-sensitive grammar

A grammar that recognizes wordforms that start with “re”, “de” or “dis”

The constraint $\langle \$Verb = :V \rangle$ checks that the following letters constitute a valid verb.



There are three types of constraints

- The operator “=” tests that the two sides are identical strings, *e.g.*:

<\$X=\$Y> tests that ALU \$X is identical to ALU \$Y

<\$N\$Nb="plural"> tests that ALU \$N number is “plural”

<\$DET\$Nb=\$N\$Nb> tests that ALUs \$DET and \$N agree in number

- Matching:

<\$X=:N> tests that \$X is a noun (<N>)

<\$Y=:have> tests that \$Y is a form of have (<have>)

<\$V=:V+Trans+F+3+s> test that \$V is a transitive verb conjugated in the Present, third person, singular (<V+Trans+F+3+s>)

- Existence:

<\$NE> tests that variable \$NE has a non-empty value, *i.e.*, the corresponding ALU occurs in the matching sequence.



CONGRATULATIONS



You know how to construct context-sensitive grammars to describe phrase structures and agreements between their components

