

2. Construire des grammaires



Les grammaires

- Sauvegarder les requêtes dans un fichier
- Nommer les requêtes :

Verbes = <décéder> | <mourir> | <périr> ;

Noms = décès | <mort> ;

Adjectifs = <fatal> | <mortel> ;

- Combiner les requêtes : dans un fichier, la requête principale a le nom **Main** :

Main = :Verbes | :Noms | :Adjectifs ;

File > New >
Grammar >
Syntax

•copier/coller
chaque requête

•nommer
chaque requête

•créer une
requête
principale **Main**

```
# Exemple de grammaire algébrique avec règles au
# et symboles lexicaux

# Voici un ensemble de verbes :

Verbes = <décéder> | <mourir> | <périr> ;

# Ensemble de noms :

Noms = <décès> | <mort> ;

# Ensemble d'adjectifs :

Adjectifs = <fatal> | <mortel> | <périssable> ;

# Voici la règle principale :

Main = :Verbes | :Noms | :Adjectifs ;
```

File Edit Lab Project Windows Info CONCORDANCE

_La femme de trente ans.not

1 / 926 TUs

Characters Tokens Digrams Annotations Unknowns

Language is "French" Text Delimiter is: \ Text contains 926 Tokens 81359 tokens including 40000 words

Show Text Annotation Structure

Au commencement du mois d'avril 1813, il y eut un dimanche dont la matinée promettait un de ces beaux jours où les Parisiens voient pour la première fois de l'année leurs pavés sans boue et leur ciel sans nuages. Avant midi, un cabriolet à pompe attelé de deux chevaux fringants déboucha dans la rue de Rivoli par la rue Castiglione, et s'arrêta derrière plusieurs équipages stationnés à la grille nouvellement ouverte au milieu de la terrasse des Feuillants. Cette leste voiture était conduite par un homme en apparence soucieux et maladif; des cheveux grisonnants couvraient à peine son crâne jaune et le faisaient vieux avant le temps: il jeta les rênes au laquais à cheval qui suivait sa voiture, et descendit pour

Locate a pattern in _La femme de trente ans

Pattern is:

- a string of characters:
- a PERL regular expression:
- a NooJ regular expression:
- a NooJ grammar:

C:\Users\Max\Documents\NooJ\fr\Syntactic Analysis_mort.nog Set

Index

- Shortest matches
- Longest matches
- All matches

Limitation

- Only: 100 matches
- All matches
- 1 example per match

Reset Concordance

N o o J

Concordance for Text _La femme de trente ans.not

Reset Display: 15 characters before, and 15 after. Display: Matches Outputs word forms

Before	Seq.	After
he était le treizième de l'année 1813 . Le surlendemain, Napoléon partait pour cette	fatale	campagne pendant laquelle il allait perdre successivement Be
être foulée. - Restons, mon père. D'ici je puis encore apercevoir l'Empereur ; s'il	périssait	pendant la campagne, je ne l'aurais jamais vu." Le père tres
Oh ! mon Dieu, oui. A Wagram au milieu du feu, à la Moscowa parmi les	morts	, il est toujours tranquille comme Baptiste lui !" Cette répons
oi ? répondit le vieillard. - Vous allez donc encore m'affliger en me parlant de votre	mort	. J'étais si gaie ! Voulez-vous bien chasser vos vilaines idées :
de rougir. Ah ! j'espérais te voir fidèle à ton vieux père jusqu'à sa	mort	, j'espérais te conserver près de moi heureuse et brillante ! t'a
à sa solitude et à sa tristesse dans le chagrin que lui avait causé la	mort	de son père, de qui elle portait encore le deuil. Au bout de h
distraction, ce jeune Anglais s'est mis à étudier sa maladie, que l'on croyait	mortelle	. Insensiblement, il a pris goût à l'anatomie, à la médecine; il s
plus les femmes qu'ils ne savent aimer; ils croient que d'aller à la	mort	le lendemain les dispense d'avoir, la veille, des égards et des

Query 127/127

7 sec Cancel

Grammaire du futur

Morpho = <V+F> ;

Proche = <aller+PR> <ADV>* <V+INF> ;

AdvFut = demain |

dans <NB> (minutes | heures | jours | semaines |
mois | ans) |

:NJour prochain ;

NJour = lundi | mardi | mercredi | jeudi |
vendredi | samedi | dimanche ;

Main = :Morpho | :Proche | :AdvFut ;

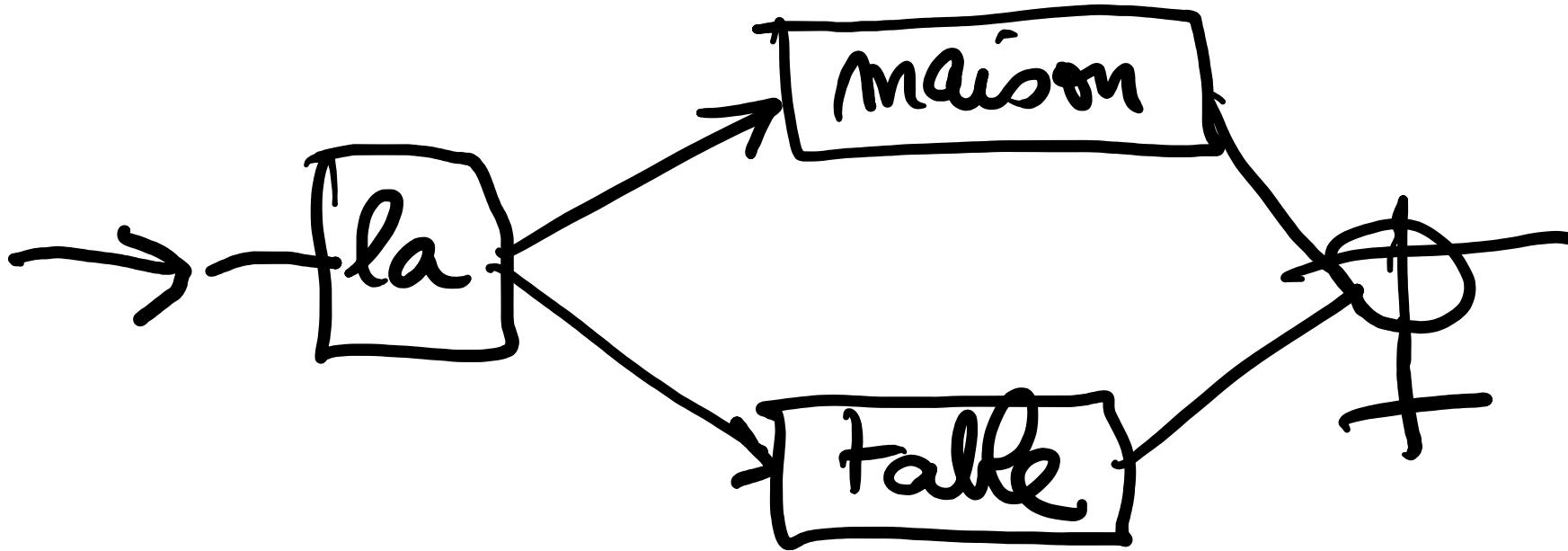
Editeur graphique de grammaires

- Une grammaire peut être représentée par un **graphe**
- Un graphe est un ensemble de **nœuds**
- Deux nœuds spéciaux : le **nœud initial** et le **nœud terminal**
- Les nœuds sont **étiquetés** par des expressions régulières (sans cycle)
- les nœuds peuvent être **connectés**

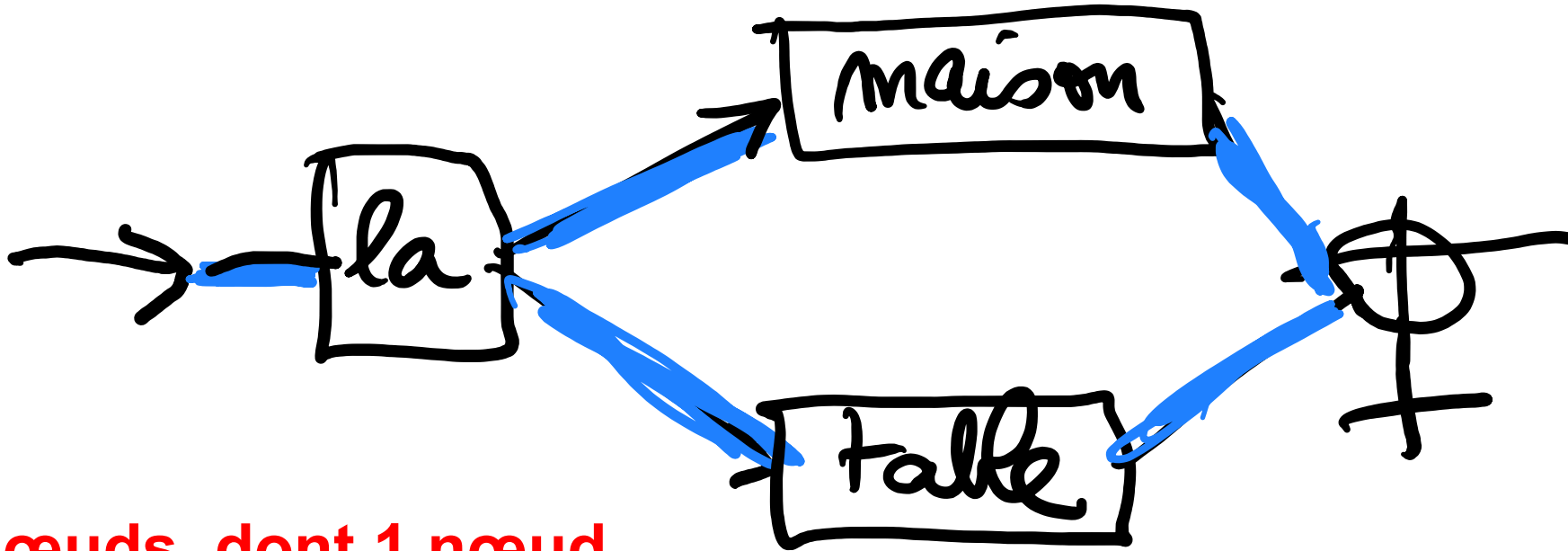
Grammaires graphiques

- Un graphe est un ensemble de nœuds
- Un graphe comporte un nœud initial, et un nœud terminal
- Les nœuds sont étiquetés par des expressions
- Les nœuds peuvent être connectés ou non ; les connections sont orientées

Un graphe



Un graphe



**5 nœuds, dont 1 nœud
initial et 1 nœud
terminal**

5 connections

Editeur graphique de grammaires

- Ctrl-Cliquer pour créer un nœud et l'éditer
- Ctrl-Cliquer sur un nœud pour l'éditer
- Ctrl-Entrée pour valider l'édition d'un nœud
- **ATTENTION : ne pas double-cliquer !**
- Cliquer sur deux nœuds pour les connecter
- **ATTENTION : les connections sont orientées !**
- Cliquer sur deux nœuds pour les déconnecter

- Créer un graphe pour : *la (maison/table)*
- Ajouter *petite, une*

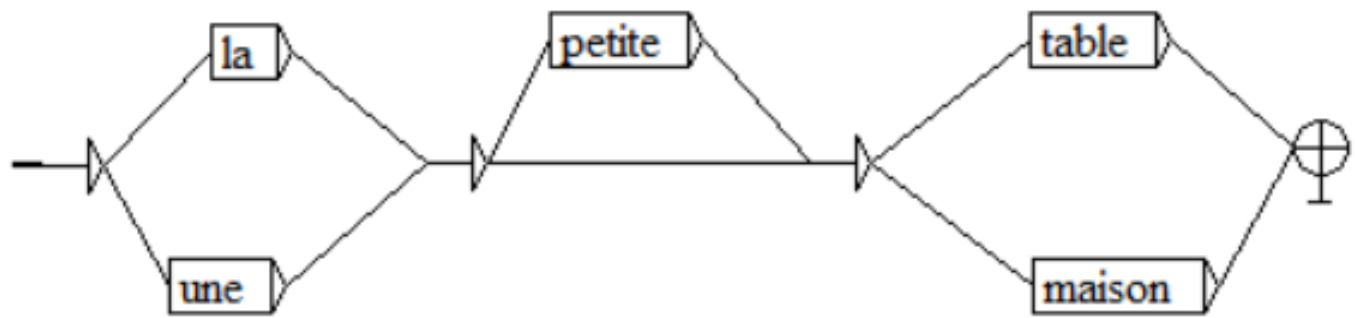
NooJ - [C:\Users\Max\Documents\NooJ\fr\Syntactic Analysis\zozo1423.nog]



File Edit Lab Project Windows Info GRAMMAR



French/French syntactic grammar.



8.7 sec

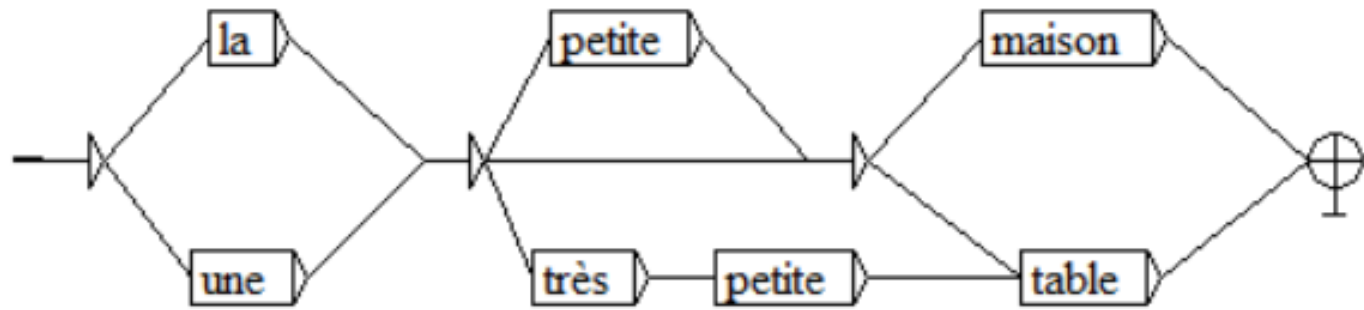
Cancel

- Créer un graphe pour : *la (maison/table)*
- Ajouter *petite, une*
- On veut en plus *très* dans "la très petite table"
mais pas "la très petite maison"

NooJ - [C:\Users\Max\Documents\NooJ\fr\Syntactic Analysis\zozo1424.nog]

File Edit Lab Project Windows Info GRAMMAR

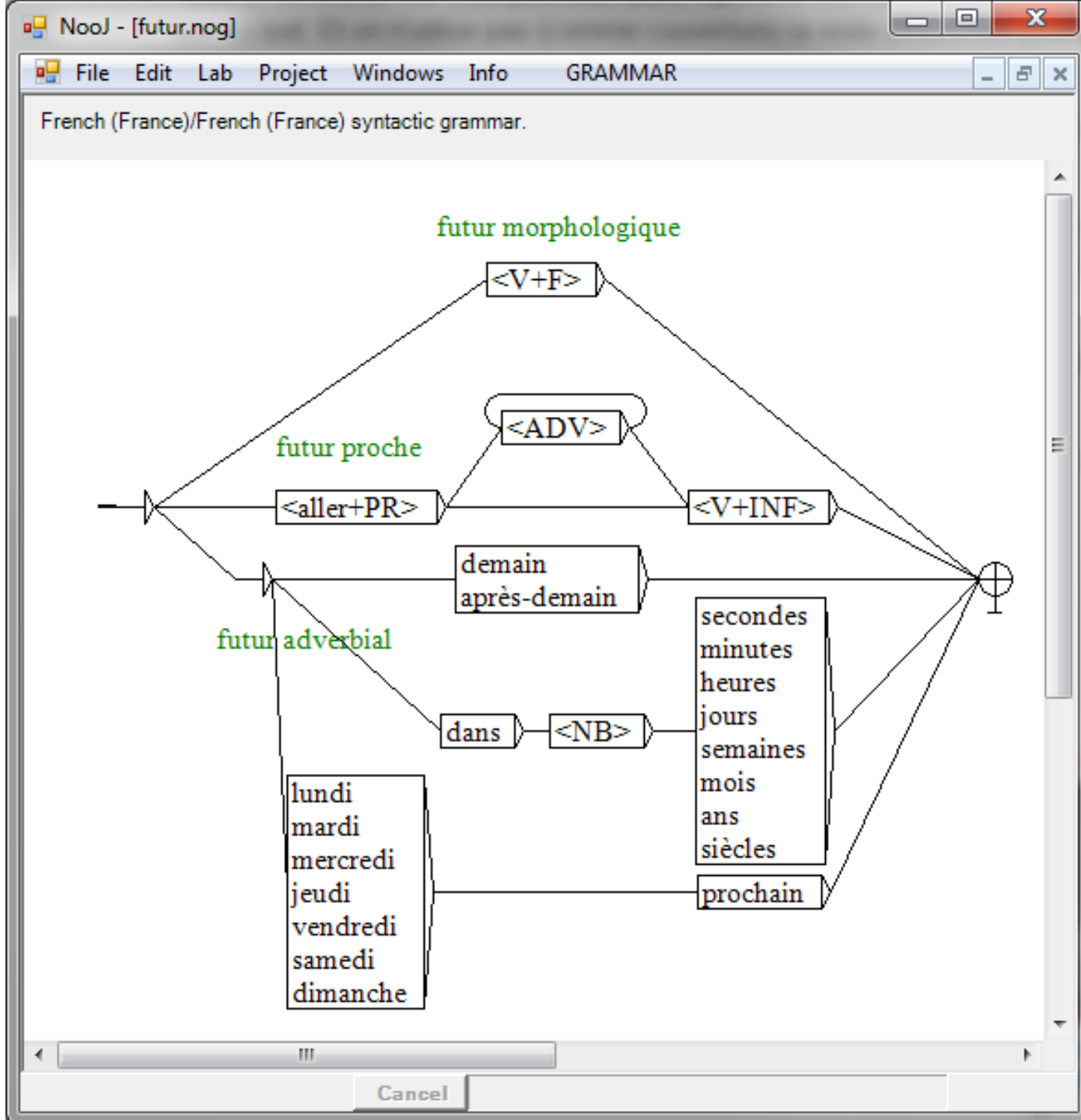
French/French syntactic grammar.



8.7 sec

Cancel

Les graphes



Exercice : grammaire locale

- Construire la grammaire qui reconnaît toutes les séquences de mots entre *il* et *donnera*
- Exemples de séquences à reconnaître :
 - Il donnera, il ne donnera, il me le donnera,
il le leur donnera, il se donnera,
il y donnera, il t'en donnera, il n'en donnera
- Contre-exemples :
 - *il lui le donnera, *il le nous donnera,
*il m' le donnera, il l'y nous donnera
*il te en donnera, *il ne en donnera

French/French syntactic grammar.

Window Snip

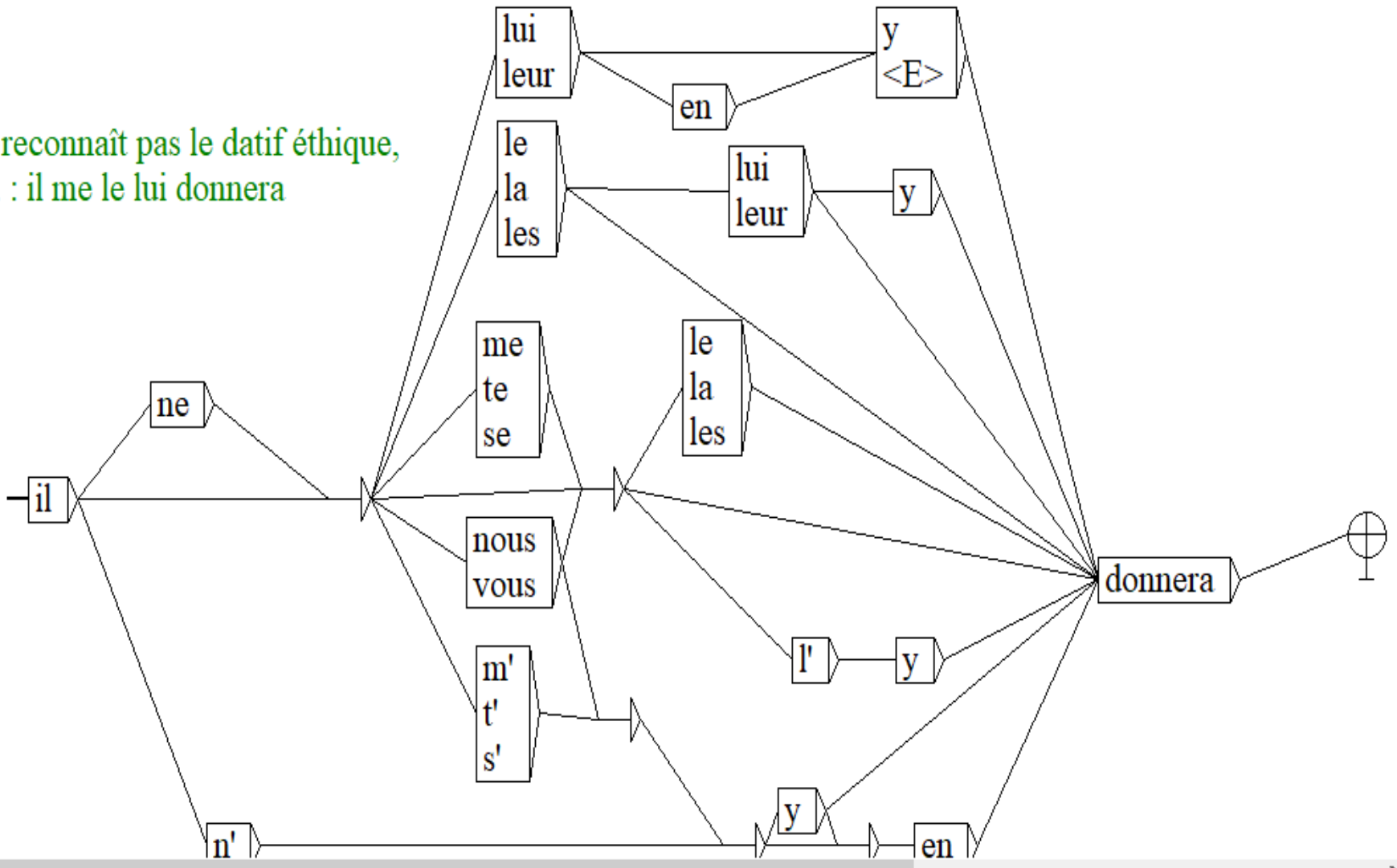
il

donnera

Cancel

French/French syntactic grammar.

ne reconnaît pas le datif éthique,
ex. : il me le lui donnera



Grammaire des dates

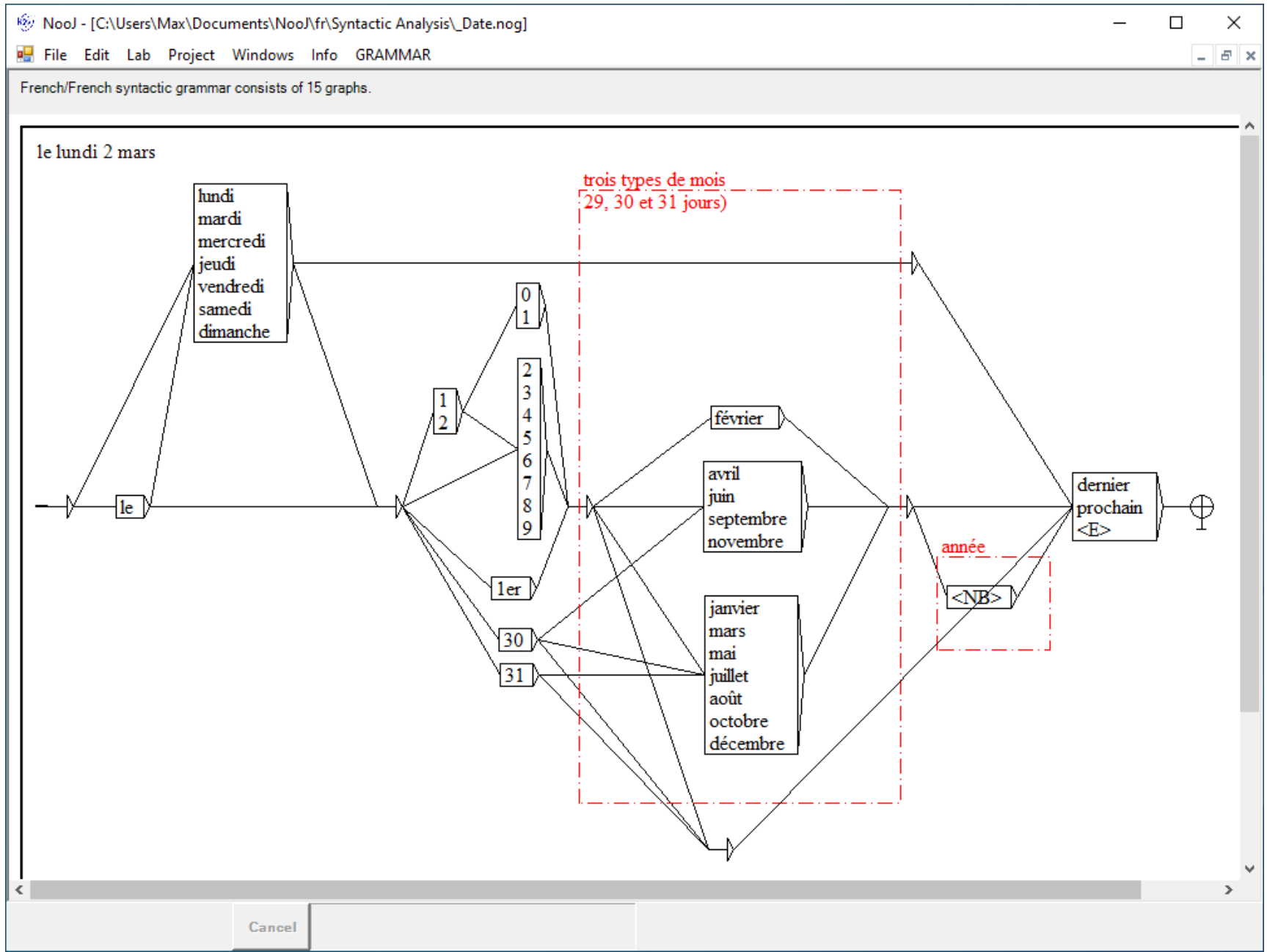
- Construire la grammaire des dates, ex. : *jeudi 18 novembre*

*le 10, mardi 1er juin, le 23 novembre,
lundi 17, jeudi, le samedi 31 janvier*

**le 0 janvier, *le 1 mai, *le septembre, *le 30
février, *jeudi 31 avril,*

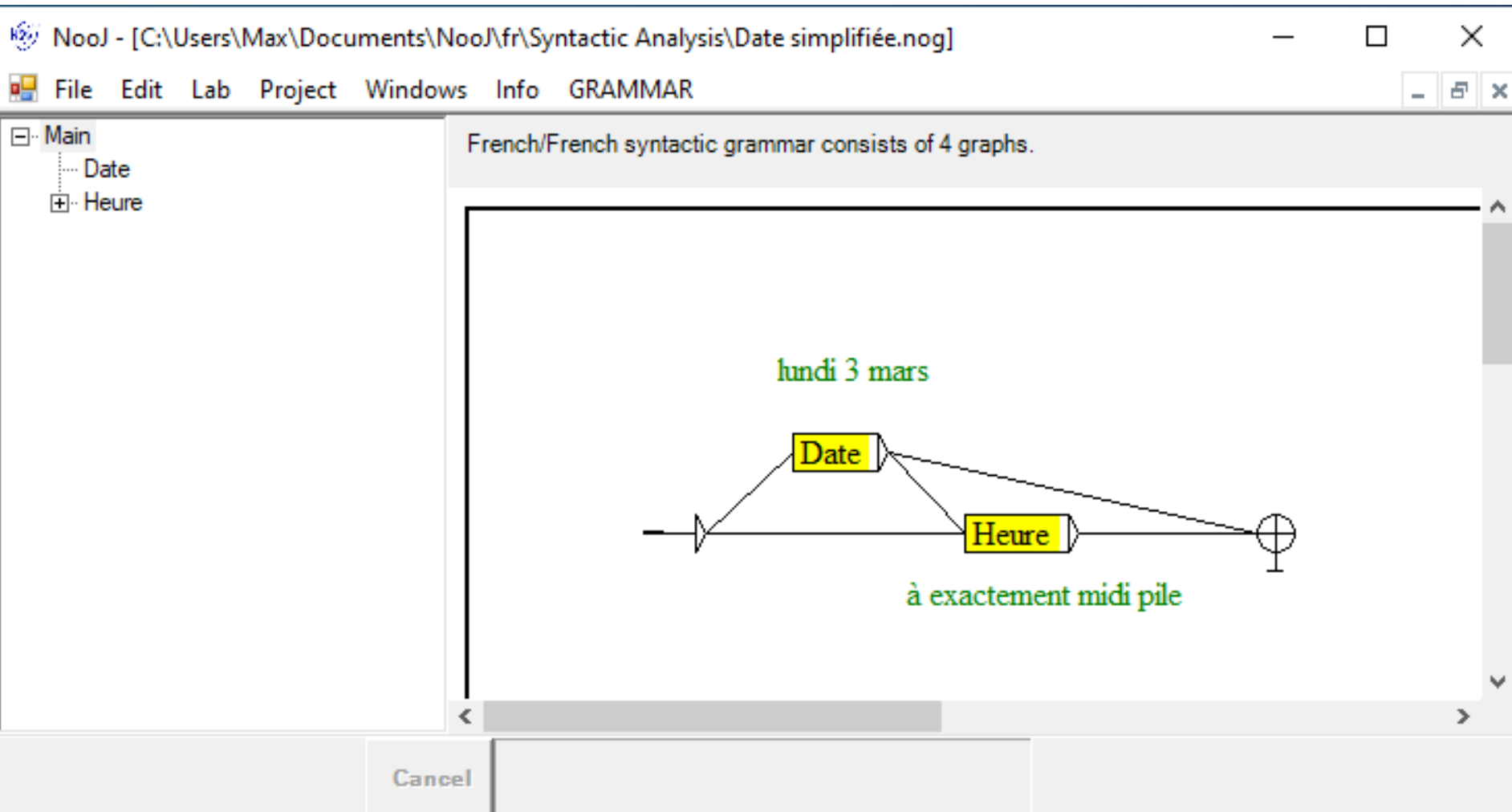
**vendredi le 25, *samedi juin, *mars, *12 mai*

Comment exprimer une date ?



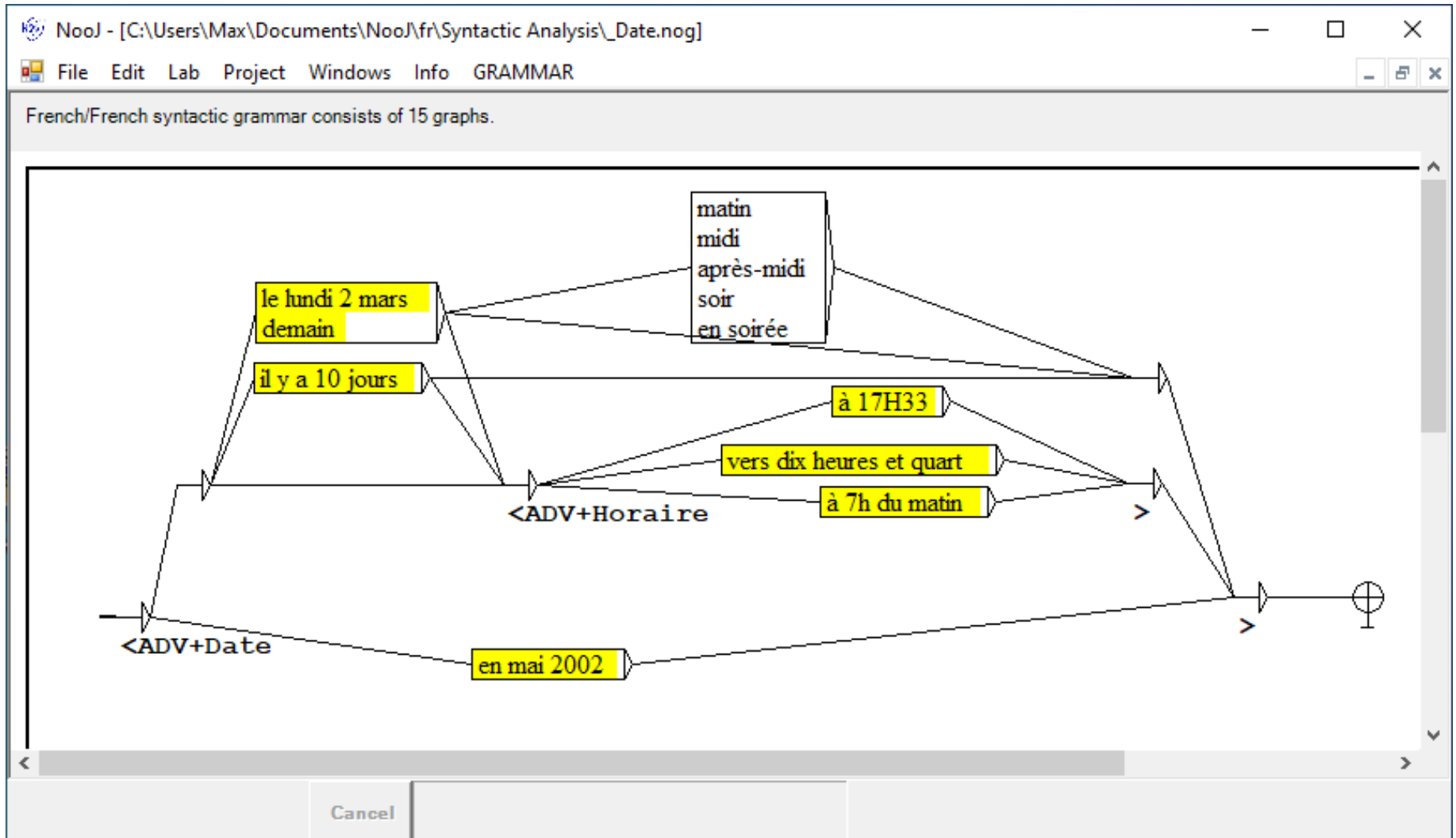
Les grammaires structurées

Quatre graphes : **Main**, **Date**, **Heure** et **moins le quart**

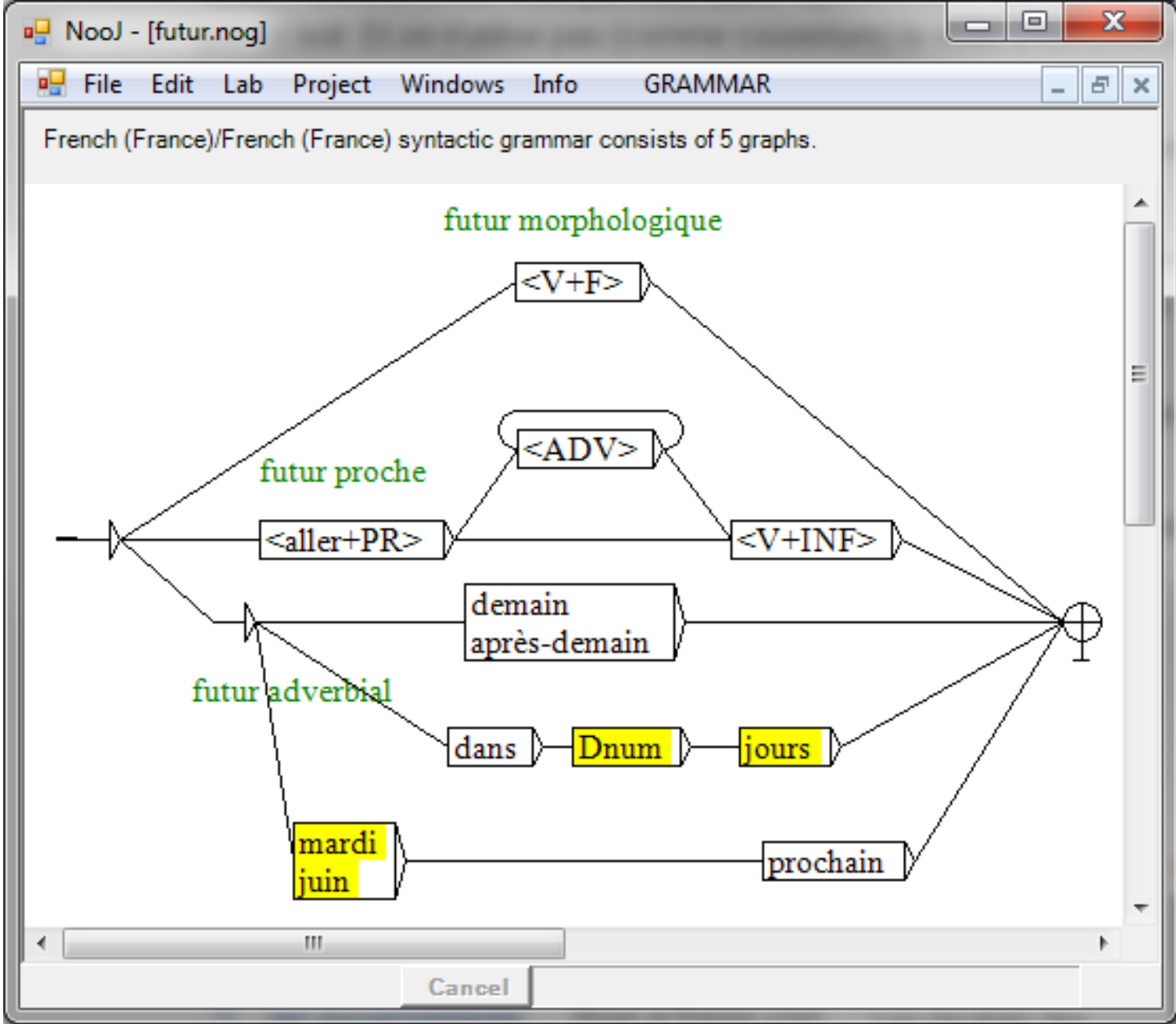


Les grammaires structurées (15 graphes)

- Show Structure
- Grammaire des nombres



Les grammaires structurées



Chaque grammaire a un contrat

NooJ - [[Modified] C:\Users\Max\Documents\NooJ\fr\Syntactic Analysis_Date.nog]

File Edit Lab Project Windows Info GRAMMAR

French/French syntactic grammar consists of 15 graphs.

le lundi 2 mars
demain

il y a 10 jours

à 17H33

vers dix heures et quart

à 7h du matin

matin
midi
après-midi
soir
en soirée

<ADV+Date

<ADV+Horaire

Contract

Check

```
# NooJ
# Enter examples, *counter-examples and #comments
#
le 1er janvier
*le 1 février

le 2 février
```

Cancel

Déboguer une grammaire

vert = séquence parfaitement reconnue

NooJ - [[Modified] C:\Users\Max\Documents\Noo\fr\Syntactic Analysis_Date.nog]

File Edit Lab Project Windows Info GRAMMAR

French/French syntactic grammar consists of 15 graphs.

le lundi 2 mars
demain

à 17H33
vers dix heures et quart
à 7h du matin

en mai 2002

matin
midi
après-midi
soir
en soirée

<ADV+Horaire>

<ADV+Date>

Enter expression:

1 analysis. Click a solution below to display the corresponding path:

Perfect Match Partial Match Failure

Paths

("Main" ("le lundi 2 mars" lundi 3 mars à midi pile) ("vers dix heures et quart" ("moins le qua

Cancel

Déboguer une grammaire

bleu = séquence partiellement reconnue

NooJ - [[Modified] C:\Users\Max\Documents\NooJ\fr\Syntactic Analysis_Date.nog]

File Edit Lab Project Windows Info GRAMMAR

French/French syntactic grammar consists of 15 graphs.

Enter expression:

1 analysis. Click a solution below to display the corresponding path: Perfect Match Partial Match Failure

Paths	Outputs
("Main" ("le lundi 2 mars" lundi 3 mars à Paris))	<ADV+Date>

Déboguer une grammaire

rouge = séquence pas du tout reconnue

NooJ - [[Modified] C:\Users\Max\Documents\NooJ\fr\Syntactic Analysis_Date.nog]

File Edit Lab Project Windows Info GRAMMAR

French/French syntactic grammar consists of 15 graphs.

Enter expression: Debug

12 analyses. Click a solution below to display the corresponding path:

Perfect Match Partial Match Failure

Paths	Outputs
("Main" ("le lundi 2 mars" le	
("Main" ("le lundi 2 mars" le	
("Main" ("le lundi 2 mars" le	
("Main" ("le lundi 2 mars" le	
("Main" ("le lundi 2 mars" le	

Cancel

Formation NooJ

Maîtriser NooJ pour :

- construire ses propres dictionnaires, inventer ses propres codes, associer les entrées du dictionnaire à des formes (fléchies, dérivées, abrégées, synonymes, etc.)
- construire des grammaires régulières, hors-contexte, contextuelles et non-restreintes
- annoter des textes, exporter les textes annotés en XML, importer des fichiers XML
- Lever les ambiguïtés (supprimer des annotations)
- Analyses statistiques : test d'indépendance, analyse factorielle, cooccurrences, évolutions
- utiliser des dictionnaires et grammaires multilingues pour construire des applications de traduction automatique
- construire des grammaires transformationnelles pour produire des paraphrases ou résumés automatiquement

Merci pour votre participation !

- N'hésitez pas à me contacter :
max.silberztein@gmail.com
- Il y a plein de gens sympas dans la communauté NooJ
- Des ateliers de formation sont organisés régulièrement
- Venez présenter votre projet à la conférence internationale NooJ, à l'INALCO, 11-13 juin 2025