Xerox Incremental Parser

XIP French Grammar

User's Guide

XEROX®

Research Centre Europe

Authored by:

Salah Ait-Mokhtar

XEROX RESEARCH CENTRE EUROPE

6 CHEMIN DE MAUPERTUIS
38240 MEYLAN
FRANCE

Table of Contents

Introduction

One aim of this document is to explain category and chunk names that are most widely used in the XIP French Grammar.

Chunks are basic syntactic domains that are computed by the French grammar and that serve as a basis for the dependency extraction.

Another aim of this document is to list the dependencies that are extracted by the XIP French Grammar. A XIP dependency is a n-ary relation (unary, binary) that binds one or two nodes together from the chunk tree. Usually, dependencies apply on lexical nodes. Each dependency is named with a specific label. A dependency label denotes the grammatical relation that has been computed between the nodes that have been bound together.

1 Categories & Features

A category in a XIP grammar must be declared as a set of features (attribute-value pairs).

The PoS (part-of-speech) features are distinguished from other features.

Lexical categories and feature names do not display on screen in the default output mode. However, in the XIP API all the linguistic information (e.g., categories and features) is available within the XipResult object.

In the default output of a XIP grammar, **category and feature names do not appear**. If you want to have them displayed, use the "-a" option (see the XIP User Manual.) The list of the main categories and features used in the XIP French Grammar is the following:

+Abr	Abreviation
+Acc	Accusative
+Adj	Adjective
+ADJ2_INV	Invariable premodifier adjective
+ADJ2_PL	Plural adjective
+ADJ2_SG	Singular premodifier adjective
+ADJ_INV	Invariable adjective
+ADJ_PL	Plural adjective
+ADJ_SG	Singular adjective
+ADV	Adverb
+AdvInt	Interrogative adverb
+Busi	Business/industry organisation
+Card	Cardinal
+City	City
+CM	Comma
+COMME	Specific tag for the word "comme"
+Cond	Conditional mood
+ConjComp	Comparative conjunction
+CONJQUE	Specific tag for the word "que"
+CONN	Connector
+COORD	Coordination
+CoordAdv	Adverbial coordination
+Country	Country
+Dat	Date
+Day	Day
+Def	Definite
+Dem	Demonstrative
+Det	Determiner
+DET_PL	Plural determiner

+DET_SG	Singular determiner
+Fem	Feminine
+Fut	Future tense
+Gen	Genitive
+Geo	Geographical entity
+Guessed	Guessed word (for unknown words)
+HYPH	Hyphen
+Imp	Imperative mood
+IMPERSO	Impersonal
+Indef	Indefinite
+IndI	Indicative imperfect
+IndP	Indicative mood and present tense
+Inf	Infinitive
+INITIAL	Initial (for names)
+Int	Interrogative
+Interj	Interjection
+Loc	Location (for pronouns)
+Location	Location
+Masc	Masculine
+Measure	Measure
+MISC	Miscellaneous (other)
+Month	Month
+NEG	Negation
+Nom	Nominative case
+Noun	Noun
+NOUN_INV	Invariable noun
+NOUN_PL	Plural noun
+NOUN_SG	Singular noun
+NUM	Numeral
+Ord	Ordinal
+Org	Organisation
+P1	First person
+P2	Second person
+P3	Third person
+PAP_INV	Invariable Past participle
+PAP_PL	Plural Past participle
+PAP_SG	Singular Past participle
+PaPrt	Present participle
+PC	Clitic pronoun
+PL	Plural
+Poss	Possesive
+PP1P	Plural possessive pronoun

+PP1S	Singular possessive pronoun
+PP2P	Plural possessive pronoun
+PP2S	Singular possessive pronoun
+PP3P	Plural possessive pronoun
+PP3S	Singular possessive pronoun
+PREFIX	Prefix
+PREP	Preposition
+PREP_A	Preposition “à”
+PREP_DE	Preposition “de”
+PREP_EN	Preposition “en”
+PRON	Pronoun
+PRON_P1P2	First or second person pronoun
+Proper	Proper noun
+PUNCT	Punctuation
+Quant	Quantifier
+Refl	Reflexive
+Region	Region
+Rel	Relative pronoun
+SENT	End of sentence
+SG	Singular
+SubjI	Imperfect subjunctive mood
+SubjP	Present subjunctive mood
+Time	Time
+Tit	Title
+VAUX_INF	Infinitive auxiliary
+VAUX_P1P2	First or second person auxiliary
+VAUX_P3PL	Plural 3rd person auxiliary
+VAUX_P3SG	Singular 3rd person auxiliary
+VAUX_PAP	Past participle auxiliary
+VAUX_PRP	Present participle auxiliary
+Verb	Verb
+VERB_INF	Infinitive verb
+VERB_P1P2	First or second person verb
+VERB_P3PL	Plural 3rd person verb
+VERB_P3SG	Singular 3rd person verb
+VERB_PRP	Present participle verb

2 Chunk Nomenclature

Chunks are basic syntactic domains that are computed by the French grammar and that serve as a basis for the dependency extraction.

Here follows the list of chunks used in the XIP French Grammar.

IV	infinitive verbal chunk
FV	finite verbal chunk
GV	gerund verbal chunk
AP	adjectival chunk
NP	nominal chunk
PP	prepositional chunk
BG	beginning of a clause (sentential) chunk (e.g. conjunctions that introduce an embedded clause)
INS	inserted chunk (e.g. comment clauses, parenthesis)
SC	sentential chunk (from the beginning of a clause until the first finite verb)

3 Dependency Nomenclature

A XIP dependency is a n-ary relation (unary, binary) that binds one or two nodes together from the chunk tree. Usually, dependencies apply on lexical nodes. Each dependency is named with a specific label. A dependency label denotes the grammatical relation that has been computed between the nodes that have been bound together. In a binary dependency relation, the first element is the head, and the second is the dependent.

List of dependencies

Below is the list of dependencies extracted by the French XIP parser. The dependency features (starting with the underscore character) are explained in the next section.

ADJMOD

This dependency attaches the modifier of an adjective to the adjective itself.

Example
Des moyens supérieurs à ceux du maire.

ADJMOD(supérieurs,ceux)

Impossible en fait de ne pas envisager de négociations.

ADJMOD(impossible,envisager)

AUXIL

This dependency attaches an auxiliary to the verb it modifies.

Example
Ils l'ont vu.

AUXIL(vu,ont)

CONNECT

This dependency links the verb of a finite clause to the grammatical word that introduces the clause.

Example
Il semble peu probable que les députés votent le projet.

CONNECT(votent,que)

La société qui lance l'OPA.

CONNECT_REL(lance,qui)

COORDITEMS

This binary relation links coordinated elements.

Example
Luc et Pierre participent au jeu.

COORDITEMS(Luc,Pierre)

COREF
This dependency attaches a (relative) pronoun to its antecedent. In the current version of the parser, only subject relative pronouns are handled.

Example
L'homme qui a libéré l'Europe centrale.

COREF_REL(homme,qui)

DEEPOBJ
This dependency attaches a deep object to the verb.

Examples
Les chefs sont conviés à une réunion.

DEEPOBJ(conviés,chefs)

Les chefs viennent d'être invités à une réunion.

DEEPOBJ(invités,chefs)

DEEPSUBJ
This dependency attaches a deep subject to the verb.

Examples
Certains pays sont tentés par un statut de neutralité.

DEEPSUBJ(tentés,statut)

DETERM
This binary relation links a nominal head and a determiner

Examples
Ces trois filles

DETERM(filles,ces)

Sur la table

DETERM(table,sur)

NEGAT
This unary dependency marks negated verbs.

Example
Cette méthode ne suffit plus.

NEGAT(suffit)

NMOD
This dependency attaches a modifier to the noun it modifies.

Examples
Elle a conduit au paroxysme actuel.

NMOD(paroxysme,actuel)

Par de savantes manoeuvre.

NMOD(manoeuvres,savantes)

Pour répondre à des situations de crise.

NMOD(situations,crise)

OBJ
This dependency attaches a direct object to its verb.

Examples
Ils a écarté ses adversaries.

OBJ(écarté,adversaries)

Sans voir que cela exigeait de la vigilence.

OBJ_PROPQUE(voir,exigeait)

Il souhaite investir.

OBJ(souhaite,investir)

PREPOBJ
This dependency attaches a preposition to the noun or the verb it precedes.

Example
Les acteurs de l'époque intervenaient souvent.

PREPOBJ(époque,de)

Ils essaient de ne pas intervenir.

PREPOBJ(intervenir,de)

REFLEX
This dependency attaches a reflexive pronoun to the verb.

Example
Il s'est dérobé aux questions.

REFLEX(dérobé,s')

SUBJ
This dependency attaches the surface subject to the verb, including infinitive verbs.

Example
Ils comptent le programmer fin juin.

SUBJ(comptent,Ils)
SUBJ(programmer,Ils)

SUBJCLIT
This dependency attaches a subject clitic to its verb.

Example
Cet espace est-il urbanisé ?

SUBJCLIT(urbanisé,-il)

VMOD
This dependency attaches a modifier of a verb to the verb itself. The modifier can be an indirect complement or an adjunct of the verb.

Example
Ils vont à Paris.

VMOD(vont,Paris)

Ils travaillent pour une banque.

VMOD(travaillent,banque)

Il travaillent pour pouvoir s'amuser.

VMOD(travaillent,pouvoir)

The Dependency Features

A list of suffixes are sometimes added to the dependency labels (described above) in order to give more specific information about the syntactic relation. Here follows the detailed list of the possible suffixes (features).

_COORD
This features (specific to the SUBJ and OBJ relations) means that the dependent in the current relation is coordinated with another element (which is also a dependent in another dependency occurrence having the same label than the current one.)

_DEF
This feature means that the determiner (dependent) of the DETERM dependency is a definite article.

_DEM
This feature means that the determiner (dependent) of the DETERM dependency is a demonstrative pronoun.

_IMPERSO
This feature (specific to the SUBJ relation) means the dependent (i.e. the subject) is an impersonal subject.

_INT
This feature means that the determiner (dependent) of the DETERM dependency is a interrogative pronoun.

_NUM
This feature means that the determiner (dependent) of the DETERM dependency is a numeral.

_PASSIVE
This feature means that the verbal head of the dependency is in a passive form.

_POSIT1
This feature (specific to the NMOD and VMOD relations) indicates that the dependent is attached to the nearest head preceding it.

_POSIT2
This feature (specific to the NMOD and VMOD relations) indicates that the dependent is attached to the second head candidate preceding it (less reliable than POSIT1 relations).

_POSIT3
This feature (specific to the NMOD and VMOD relations) indicates that the dependent is attached to the third head candidate preceding it (less reliable than POSIT2 relations).

_POSS
This feature means that the determiner (dependent) of the DETERM dependency is a possessive pronoun.

_PROPQUE
This feature means that the dependent is a clause introduced with the grammatical word “que”.

_QUANT
This feature means that the determiner (dependent) of the DETERM dependency is a quantifier (e.g. “tous”).

_REL
This feature means that the verbal head of the dependency is the main verb of a relative clause.

_SC
This feature means that both elements of the COORDITEMS dependency are verbal heads of clauses.

_SPRED
This feature (specific to the OBJ relation) indicates that the object of the (copula) verb is a subject predicate.

_SUBORD
This feature (specific to the VMOD relation) means that the dependent (i.e. the verb modifier) is a subordinate clause.