

FreeLing

<http://www.lsi.upc.edu/~nlp/freeling>

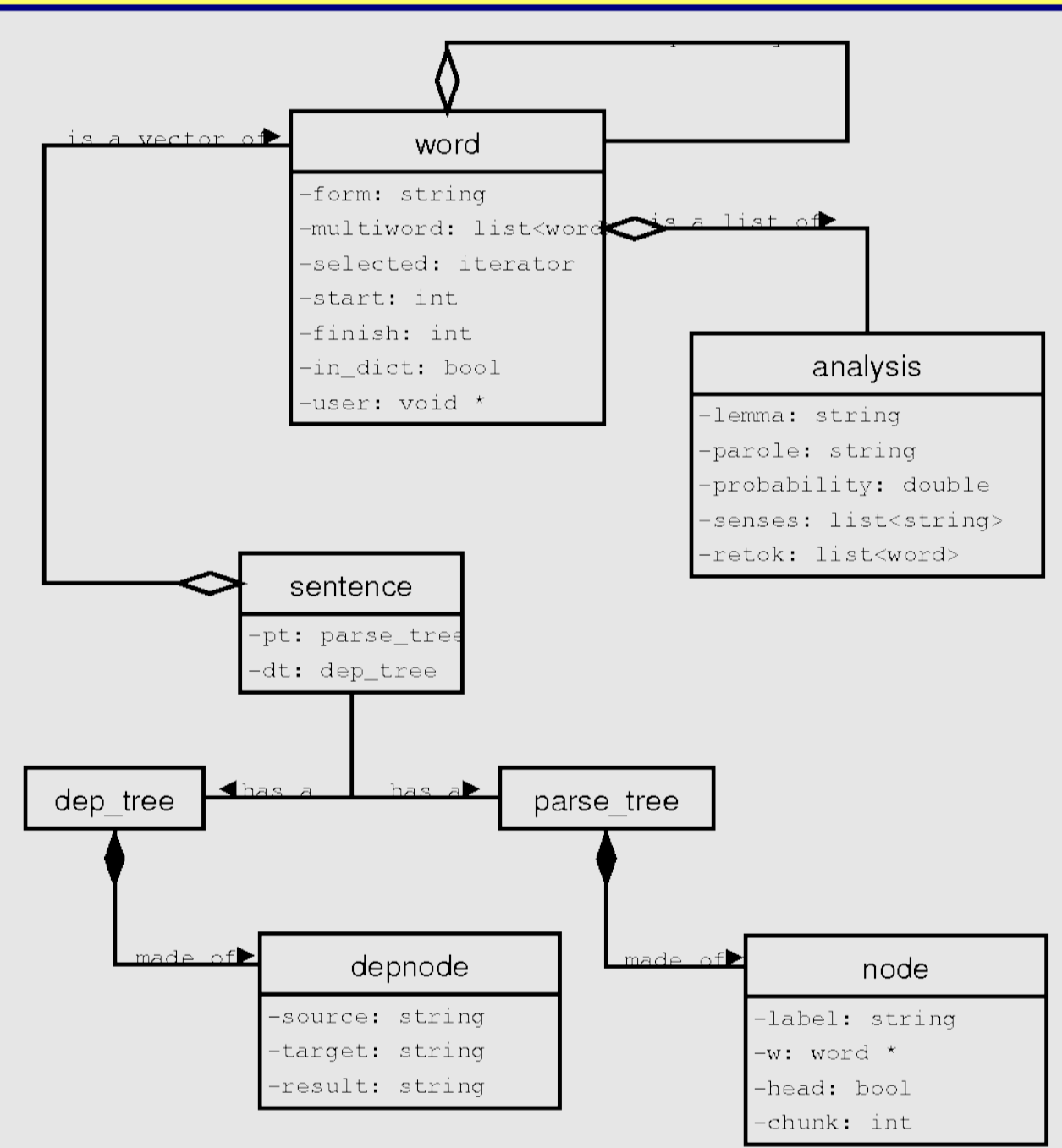
Syntactic and semantic services in an open-source NLP library

FreeLing is an open source C++ library providing language analysis services to any calling application.

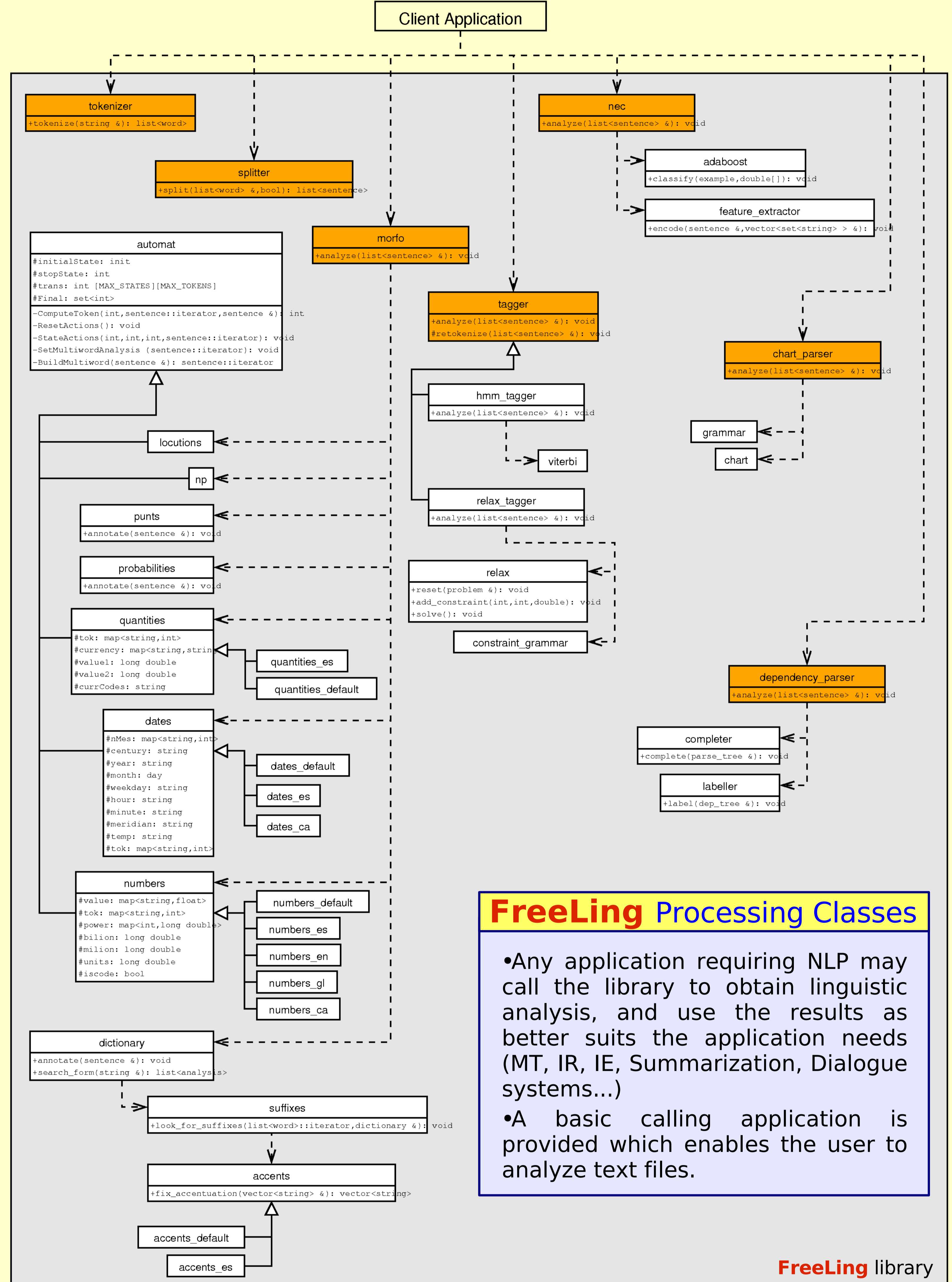
FreeLing Main features

- The library can be called from NLP applications in C++ (Java, Perl and Python are also partially supported) or be used as a standalone language analyzer via the included command-line interface.
- Spanish, Catalan, Galician, English, and Italian are covered in current version (1.4).
- Customizable behaviour via configuration files
- The library design enables easy integration of new languages.
- HMM PoS Tagging speed about 10,000 words/sec
- Provided services:
 - Tokenizing
 - Sentence splitting
 - Morphological analysis
 - Named Entity detection and classification
 - Recognition of dates, numbers, currency expressions, physical magnitudes, ratios, ...
 - PoS tagging
 - Chart-based shallow parsing.
 - Dependency parsing
 - WordNet-based semantic annotation
 - Most-frequent-sense semantic disambiguation

FreeLing Linguistic data classes



- The library returns the analysis in a set of simple linguistic classes, based on C++ STL templates.
- The calling application can customize the obtained objects to meet its needs.
- The calling application can use the data structures as received, or map the results to the desired data structure or format (XML, SVG, ...)



FreeLing Processing Classes

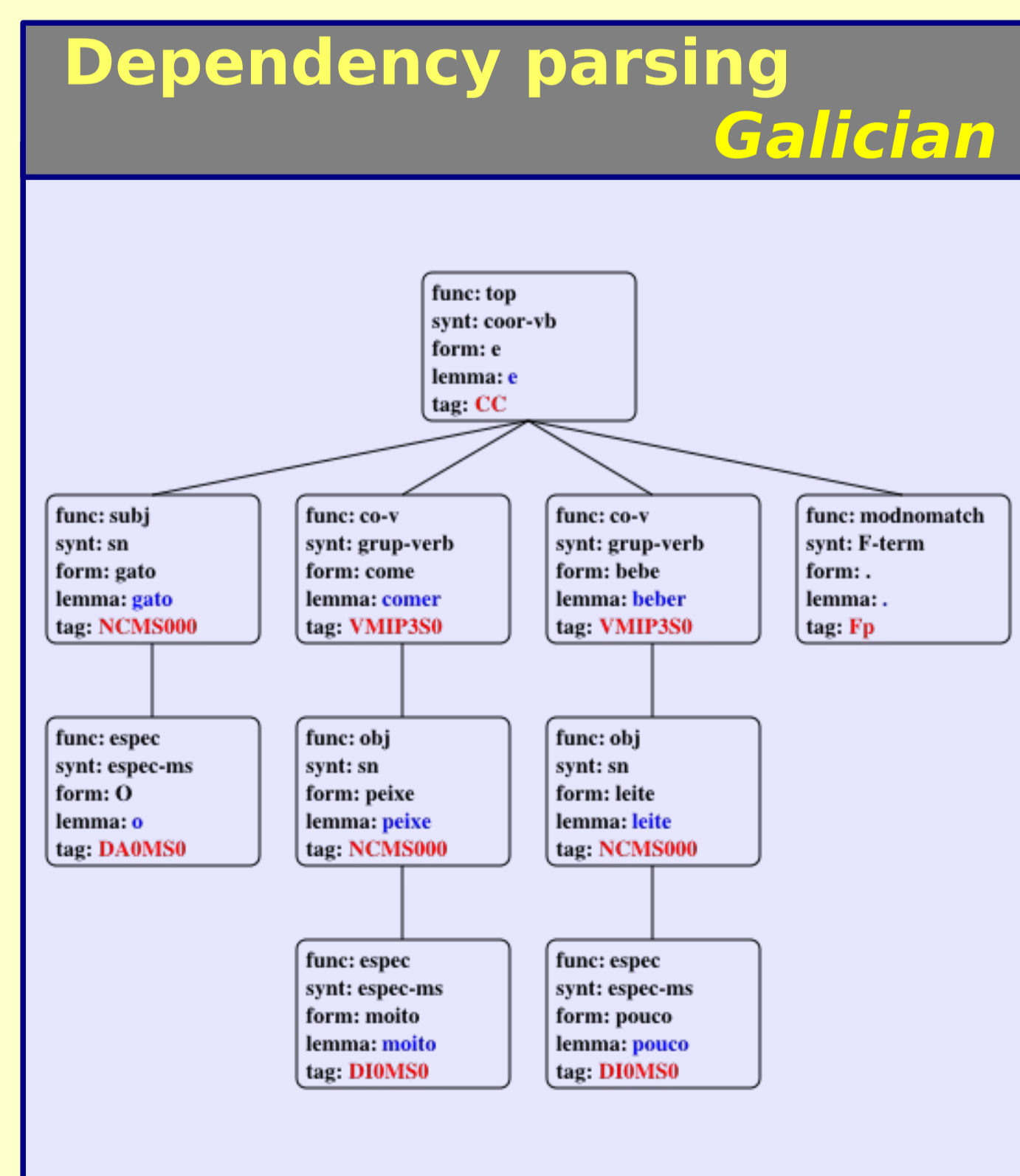
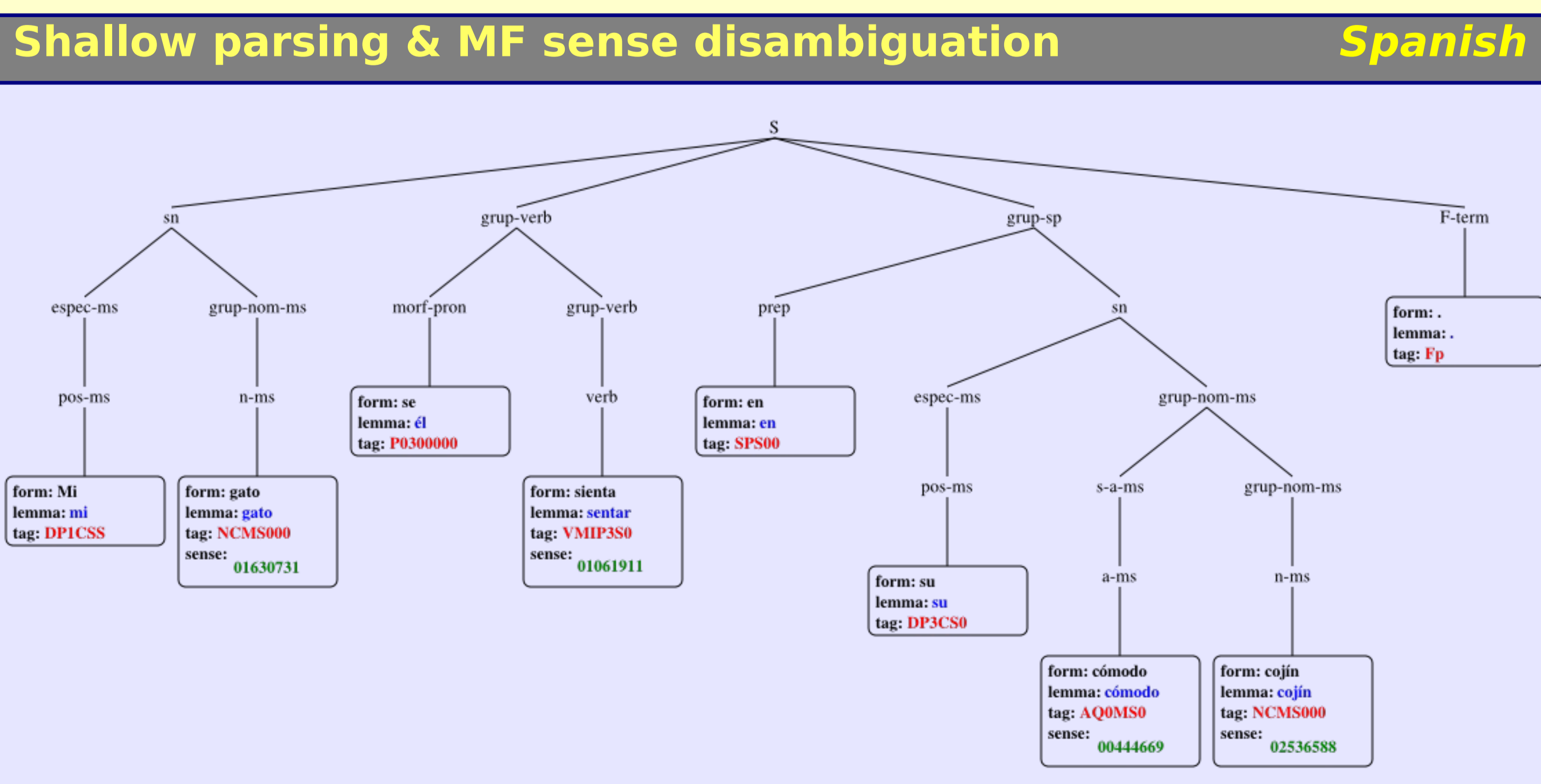
- Any application requiring NLP may call the library to obtain linguistic analysis, and use the results as better suits the application needs (MT, IR, IE, Summarization, Dialogue systems...)
- A basic calling application is provided which enables the user to analyze text files.

FreeLing library

FreeLing Output Samples

Morphological analysis							Italian
Il	mio	gatto	riposa	sulla	sua	stuoia	.
<i>il</i>	<i>mio</i>	<i>gatto</i>	<i>riposare</i>	<i>sulla</i>	<i>sua</i>	<i>stuoia</i>	<i>.</i>
DA0MS0	DP1MSS	NCMS000	VMM02S0	SPCFS	DP3FSS	NCFS000	Fp
1.00	0,91	1.00	0	1.00	0,98	1.00	1.00
	<i>mio</i>		<i>riposare</i>		<i>sua</i>		
	PX1MS0S0		VMIP3S0		PX3FS0S0		
	0,09		1		0,02		

PoS tagging & WN senses							English
My	cat	sat	on	its	cozy	mat	.
<i>My</i>	<i>cat</i>	<i>sit</i>	<i>on</i>	<i>its</i>	<i>cozy</i>	<i>mat</i>	<i>.</i>
PP\$	NN	VBD	IN	PP\$	JJ	NN	Fp
1630731	1061911				0445103	2971576	
7306044	1837041				1023508	2971832	
7143161	1352850				0421697	2971685	
2406193	1383406					3875338	
2404497	1466830					3135113	
1636523	1333288						
	1673867						
	1062571						



FreeLing is free-software, distributed under GNU Lesser General Public License (LGPL)

- The library and linguistic data are open-source, and can be freely extended, improved, customized, and redistributed.
- The library can be used as a module in commercial applications.
- Modifications to the code and data must be redistributed under the same open-source license conditions, reverting in common benefit both for research and industrial NLP community.



Centre de Tecnologies i Aplicacions del Llenguatge i la Parla
UNIVERSITAT POLITÈCNICA DE CATALUNYA