

A Yacc Grammar for the Negfinder Parser

```

/* Terminal symbols, i.e., tokens passed from the lexer to the parser, are italicized within single quotes */
sentence : clause '' | sentence clause ''';
clause : /*empty */ | clause negation | clause 'word' | clause conceptset ;

/* A sentence consists of one or more clauses. Each clause consists of one or more of the following, interspersed: a negation, a set of concepts, or other words */

/* A set of concepts consists of a list of concepts, or a list of concepts with intervening 'and' or 'or'. There may be intervening 'fillwords' - stop-words or words not recognized as concepts in UMLS */
conceptset: conceptlist | CL_or | CL_and ;
conceptlist : 'concept' | conceptlist ',' 'concept' | conceptlist 'fillword';
CL_or : 'or' | conceptlist 'or' | CL_or 'or' | CL_or concept | CL_or 'fillword';
CL_and : 'and' | conceptlist 'and' | CL_and 'and' | CL_and concept | CL_and 'fillword';

/* a negative may be either a prefix or a postfix negative. The former is of two kinds, with the negated concepts chained by AND or OR */
negative: prefixnegative | postfixnegative ;
prefixnegative: prefixnegativeand | prefixnegativeor;
of_for_igword : 'of' | 'for' | 'igword';

/* igword, for ignored word, refers to a word that is not recognized by the lexer */
prefixnegativeor : 'prefixnegor' | prefixnegativeor of_for_igword ;
prefixnegativeand : 'prefixnegand' | wordnegative 'for' | prefixnegativeand of_for_igword;
postfixnegative : 'none' | wordnegative | 'postfixneg';
wordnegative : 'negative' | wordnegative 'igword' | wordnegative 'of' ;

/* Finally, a negation may be a nonword, a nonconcept, a prefix negative optionally followed by a set of concepts, or a postfix negative optionally preceded by a set of concepts. */

/* Both nonword and nonconcept include one of the words/phrases- not, non, never, cannot, unable, unlikely to, or a pattern ending in n't (e.g., don't) followed by one or more adverbs. This pattern is followed by a concept ID pattern (for nonconcept) or a word for nonword (if it corresponds to a concept not matched to UMLS). */
negation : | 'nonword' | 'nonconcept'
| prefixnegative | prefixnegative conceptset;
| postfixnegative | conceptset postfixnegative;

```