

Prolog Tutorial 4

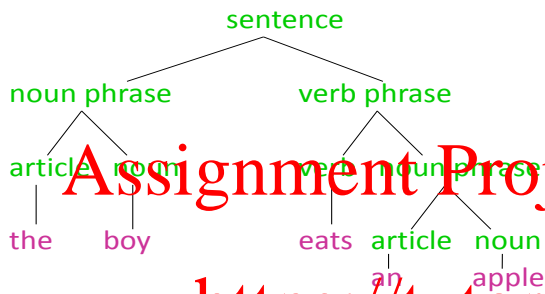
Language Processing in Prolog

Assume a grammar for very simple English:

```
sentence --> nounphrase, verb phrase
noun phrase --> article, noun
verb phrase --> verb | verb, noun phrase
```

providing, for example, the following parse tree for the sentence:

“The boy eats an apple.”



With a simple Lexicon such as:

```
article --> the | a | an
noun --> boy | apple | song
verb --> eats | sings
```

1. Write a Prolog program to check whether or not an English sentence is grammatically correct, according to the grammar. The program should also be able to generate grammatically correct sentences. For example:

“The boy eats an apple.”	is grammatically correct.
“The boy eats a eats.”	is grammatically incorrect.

Do this by

- Representing sentences as a list of words, e.g. [the, boy, eats, an, apple]
- Defining a predicate *sentence/1*, and any other auxiliary predicate you need, such that *sentence(S)* succeeds if *S* is a correct grammatical sentence
- Testing the program with your own lexicon, e.g. nouns: grass, cow, computer, girl, boy, etc.

So for example:

?- sentence([a, cow, eats, the, grass]).

Gets the answer yes.

?- sentence([girl, sings, eats]).

Gets the answer no.

?- sentence(S).

Gets all instances of S that are grammatically correct.

You will get many ridiculous sentences (like [a, grass, eats, a, cow]), because the grammar is very simple – don't worry!

2. Modify the program for recognizing and constructing noun phrases so that it requires the article to match the noun in noun phrases, i.e. if the noun starts with a vowel then the article is either 'the' or 'an', not 'a', and if the noun starts with a consonant then the article is either 'the' or 'a', not 'an'. So, for example, [a, apple] and [an, cow] should not be recognized or generated as a noun phrases.

Hint. Prolog has a built-in predicate *atom_chars/2* such that *atom_chars(W, L)* succeeds when *L* is the list of characters in the constant *W* in the order in which they appear in *W*. So *atom_chars(apple, L)* succeeds with *L=[a,p,p,l,e]*.

3. Continue the work on natural language processing by extending the grammar given so far so that your program is more *context-sensitive*:

a) Make sure the verb and the noun agree in being both singular or both plural. For example "the boys eats an apple" should not be recognised as a sentence. But "a boy eats the apples" is fine.

b) How would you modify your program to avoid strings such as "a, grass, eats, a, cow", or "a, song, chews, the, grass" being recognized as sentences?