Lesson #1: Introduction to Formal Languages

What's It All About?

The basic concepts associated formal languages, and grammars which define formal languages, are presented in this lesson.

What is a formal language?

Definition: A **formal language** is a set of strings of symbols defined over some finite vocabulary of symbols.

Examples:

1. The set of binary numbers, with leading zeros but no signs, is a language.
2. The set of integers, with neither leading zeros nor signs, is a language.
3. The set of decimal numbers, with signs but no leading zeros, is a language.
4. The set of MIDI numbers.
5. The set of tokens in the Java programming language.
6. The set of Java programs.
7. The set of words in the Scrabble dicionary.
8. The notes in JFugue notation.
9. The notes in ABC notation.

Notes on formal languages:

1. There is no mention of such things as *communication* or *meaning* in the definition of a formal language.
2. Some of the example languages are finite, some are infinite.

What is a grammar?

A **grammar** is a finite specification of a language.

There are many types of grammars, including:

1. Exhaustive enumerations
2. Natural languages
3. Phrase structure grammars

4. Recursive transtion networks

Some yes/no questions about grammars

Please answer "yes" or "no" to each question:

1. Will exhaustive enumeration work to define the language of binary numbers?
2. Will exhaustive enumeration work to define the language consisting of the words in the scrabble dictionary?
3. Is natural language well suited to precisely describing a computer programming language?
4. Are phrase structure grammars well suited to precisely describing computer programming languages? (Take a guess, even if you have no idea what a phrase structure grammar is. I bet you can blindly guess the right answer!)
5. Do you think that Post production systems constitute a kind of grammar?

Something to do ...

Consider the formal language defined over the vocabulary of symbols {LA, DI, DA} which consists of the set of all strings of these symbols beginning with LA and ending with DA.

Please write down 5 sentences in this language, each of a different length.

Some questions ...

1. What is a formal language?
2. Is the Ladida language finite or infinte?
3. What is a grammar?
4. Do you think that one can define the Ladida language by means of a natural language grammar?