N87-25885

SOLAR PREDICTION AND INTELLIGENT MACHINES

Gordon G. Johnson

Professor

Department of Mathematics

University of Houston-University Park

Houston, Texas 77004

There are two projects running concurrently, one dealing with solar prediction and the second is developing an intelligent machine.

The solar prediction program is aimed at reducing or eliminating the need to throughly understand the process previously developed and to still be able to produce a prediction. Substantial progress has been made in identifying the procedures to be coded as well as testing some of the presently coded work.

The second project involves work on developing ideas and software that should result in a machine capable of learning as well as carrying on an intelligent conversation over a wide range of topics. The underlying idea is to use primitive ideas and construct higher order ideas from these, which can then be easily related to one to another.

NASA Colleague: Robert T. Savely FM7 X4751

SOLAR PREDICTION AND INTELLIGENT MACHINES

Gordon G. Johnson
Professor
Department of Mathematics
University of Houston-University Park
Houston, Texas 77004

There are two projects being developed concurrently, the first deals with the prediction of solar activety as measured by sunspot numbers and the second project pertains to the development of an intelligent machine.

The method developed in 1978-79 for predicting monthly mean sunspot numbers[1] has been refined in order to reduce the amount of time and expertise required to produce an eight to ten year prediction of monthly mean sunspot numbers. The availability of such predictions are clearly useful for well known technical reasons such as communications and orbit determination.

The refined method was developed with the aid of the work of J. Parker and M. Boarnet that provided an environment in which to ferret out the step by step procedure that a person generating a prediction actually does, and to then capture this procedure in code, thus removing the requirement that a person throughly understand the process before attempting to generate a prediction.

Mr. J. Parker created LISP code to perform the extensive calculations required. His work adds several enhancements to the original FORTRAN code in that it allows maximun use of the available data at each step in the computations.

Mr. M. Boarnet then built an interface that provides an easily operated interactive interface as well as graphics displays, both of which were essential in the development work leading to the present procedure. The interface allows the operator to control all essential parameters that guide the calculations needed to generate the prediction.

It was found that the implementation on the SYMBOLICS machine required excessive machine time, on the order of fifty hours. To reduce this time by a substantial amount, the calculations will be done on the FLEX to gain from the advantages of parallel processing. The code is being written in **C** by K. Fields for the FLEX, who has added enhancements to the previously written code, which should result in shorter execution time on the order of four to five hours. The results of the computations will then be transferred to the SYMBOLICS for display.

When the process is fully operational on the FLEX, then the daily sunspot prediction will be tried. This would involve a thirty fold increase in data to be processed. However, it is planned that the long term monthly mean prediction would have been generated, and fine grain daily prediction will then be made, and should require only a small amount of additional time on the order of four hours.

The underlying support for this method is the anharmonic technique of A. K. Paul[2] which provides us with a tool to search for and

extract particular sine functions from the solar data.

The search is restricted to sine functions i.e., frequencies, amplitudes and phases, that have periods between six months and forty years and amplitudes greater than some preset minimum. There is no attempt to completely decompose the data set into a sum of sine functions, but rather to find approximately thirty sine functions such that the sum of those found, when subtracted from the original data set, results in a "smooth" curve that can be extrapolated. After the extrapolation is completed the subtracted sine functions are added back, resulting in the original data set as well as the predicted values.

The second project concerns the work of developing ideas and code that should result in a machine capable of learning as well as carrying on an intelligent conversation[3]. The code is being written in COMMON LISP by L. Wang on the SYMBOLICS.

Preliminary coding to find first and second level relationships has been written . The underlying notion that drives this development is that of primitive ideas. Ideas contained in the words such as toward , hold, place, time, are examples of primitive ideas. The number of words in the list of primitves is about thirty five, and it is from this short list that all the ideas are formed. The ideas form patterns that are then compared and related.

The next part to be encoded deals with analogies that should produce "new" ideas. The machine will not store results given in response to statements but each time will reconstruct anew the proper response.

If the machine does not understand a statement given to it it will inquire as to the meaning of the word or words that it does not recognize. The response to a machines inquiry will be stored and internally related to previously understood ideas. The new ideas given will then be used to resopd to the initial statement. In addition, when a statement is given to the machine the machine will then, unbeknowst to the person communicating with the machine, form a view of all the ideas that relate to the subjects in in the first statement, so that if additional statements that deal with ideas similiar to those initially given to the machine are presented to the machine it will be able to respond quickly.

There are many levels within the machine, the first consists of the ideas closely and directly related to a statement given the machine, the second level consists of all ideas directly related to content of the statement, the third level consists of all ideas related to the ideas in the second level. The higher levels are formed in a similiar fashion. The forming of the first two levels is automatic. The machines response to the initial statement is usually formed from the ideas found in the first two levels. If there is insufficient information to respond, then the machine proceeds to higher levels in its search for a response. The various ideas in the various levels are retained in preparation for expected statements related to the first statement. If a second statement that is related to the first statement is given to the machine it may cause the machine to enlarge the pool of ideas in all levels. However, if the second statement is on an entirely different topic, then the machine will form new first, and second levels.

This report is a brief description of the development, and is only the start of a much more complex involvement with machines that should result in a machine that will converse in a "natural" manner with any person.

REFERENCES

1. G. G. Johnson and S. R. Newman, "Solar Activity Prediction of Sunspot Numbers", NASA, JSC-16390 1980.

2. A. K. Paul, "Anharmonic Analysis", Mathematics of Compution, Vol.26, No. 118, April 1972.

3. T. Winograd, "Language as a Cognitive Process", Vol. 1: Syntax, Addison-Wesley Pub. Co. Reading, Mass., 1983.