Editorial

Natural Language Computations and Plant Health Concepts

DAVID R. MacKENZIE
Professor and Head, Department of Plant Pathology and Crop Physiology, Louisiana State University, Baton Rouge

I was once told that plant pathologists seem to delight in arguing terms and definitions. I took that remark as a compliment. I have often recognized that one of our profession's attractions is the excitement and challenge of dealing with complex concepts.

The profession of plant pathology has grown from a descriptive to a quantitative discipline. Throughout our history we have made major contributions to plant health. My reflections here should be taken not as criticism of our accomplishments but as genuine interest in continuing our progress.

The effort we invest in developing concepts and definitions should have better use for the advancement of our science. Our problem is that a good, solid definition of a concept like host plant resistance defies our best attempts. Plant pathologists seem to be able to assemble in small numbers and recognize desirable plants that are resistant, and we can often get general agreement on those plants that are susceptible. The description of the continuum of host plant resistance, however, has not been quantified and is further complicated by such modifiers as "vertical or horizontal" and "partial or monogenic."

I have selected host plant resistance as but one example of the many complex concepts with which we must deal. Many other concepts could have been selected. The point remains. We need concepts and definitions to communicate within our profession and with sister disciplines and to apply them to problems.

I have given a lot of thought to this predicament. Imagine my delight when I discovered the remarkable progress that has been made in computer science on natural language computations (NLC). The proponents of NLC recognize that linguistic variables—variables whose values are not numbers but are words or sentences in a natural or synthetic language—cannot be analyzed classically. They have offered us new methodologies.

The theoretical foundations of NLC are based in set theory, the so-called modern math. The methods of NLC permit manipulation of an imprecise term such as "young" with modifiers ("somewhat young") and in strings ("very young to fairly young"). Our challenge as plant pathologists is to look for the applications of this new math to our researchable problems. I give three examples. There are obviously many more.

Pest forecasting. The complex biological and environmental factors that regulate pest outbreaks are often the consequence of unquantifiable variables. Consider Wallin's expert system of severity values used for potato late blight forecasting. They are a good example of cognitive values applied to complex biology. It is likely that many more applications of NLC will be found to allow integration of concepts in host plant resistance, sanitation, epidemic development, fungicidal control, etc., for forecasting pest outbreaks.

Crop stress impact assessments. It is generally recognized that the immense cost of assembling large data bases for calculating annual crop losses is beyond our resources. I see, however, possibilities for developing far less costly "expert opinion" crop stress impact assessments that would combine linguistic descriptions and numerical records for a national report on the impact of crop stresses. The targeted variables might be descriptions on the frequency of stresses and their distribution. Analytical methods of NLC could be used to examine interactions and answer "what-if" questions.

Release of r-DNA organisms into the environment. Some individuals propose to use risk assessment for regulating the release of r-DNA organisms into the environment. Risk analysis is based on statistical and probability theory. The method seems to work well for such things as engineering problems and chemical spills—but the quantification of biotic events is far more difficult. The product-of-probabilities-of-independent-events rule has obvious flaws when applied to biology. NLC would provide far better guidance to those who need to oversee r-DNA research and to those who need to regulate the products of biotechnology.

I have avoided the term "fuzzy set analysis" used by computer scientists to describe this new methodology. Far too many people are turned off by the word "fuzzy"—perhaps because it seems to be unprofessional. I "unscientifically" tested the response of individuals to the term "fuzzy" by placing a copy of *Fuzzy Sets, Natural Language Computations, and Risk Analysis* by Kurt J. Schmucker (Computer Science Press, Inc., Rockville, MD, 1984, 192 pages) on the corner of my desk. I then gathered unsolicited comments for 2 weeks. Most reactions were negative, and many people joked about "fuzzy sets." No one asked to borrow the book.

It is truly unfortunate that computer scientists continue to use the term "fuzzy" to describe this important and exciting mathematical framework for manipulating complex or imprecise concepts. Perhaps some plant pathologist can come up with a better term. After all, we are known for devoting a lot of effort to such activities.

Anyone wishing to read more on the subject should see Schmucker's book and L. A. Zadaeh's article "A Fuzzy-Algorithmic Approach to the Definition of Complex or Imprecise Concepts" (Int. J. Man-Mach. Stud. 8:249-291, 1976).