

## Soft Computing in Bioinformatics

Jim Keller

Electrical and Computer Engineering Department  
University of Missouri-Columbia  
kellerj@missouri.edu

**Abstract:** This tutorial will go over the basics of creating and using similarities between gene products, i.e., proteins or mRNA. We discuss the various categories of similarity measures beginning with the traditional BLAST scores on gene product sequences and culminating with recent research results in the application of fuzzy set theory to this problem. Clustering, knowledge discovery and applications to microarray experiments will be stressed. The following is a topical outline:

- I. Background
  - Genes and Gene Products: Sequences, Structure, Microarrays (expression, hypermethylation)
  - Gene Ontology
  - Other taxonomies
- II. Gene Product Similarity Measures
  - BLAST
  - Smith-Waterman
  - GO-based measures: Jaccard, Cosine, Dice, Fuzzy measures, Choquet Integrals
  - Domain and Motif measures
- III. Visualization and Clustering
  - Hierarchical clustering
  - Visual Assessment of cluster Tendency
  - NERFCM
- IV. Knowledge Discovery
- V. Function Summarization

**About the speaker:** James M. Keller received the Ph.D. in Mathematics in 1978. He currently holds the rank of Professor in the Electrical and Computer Engineering Department at the University of Missouri-Columbia. He is also the R. L. Tatum Research Professor in the College of Engineering. His research interests center on computational intelligence: fuzzy set theory and fuzzy logic, neural networks, and evolutionary computation with a focus on problems in computer vision, pattern recognition, and information fusion including bioinformatics, spatial reasoning in robotics, sensor and information analysis in technology for eldercare, and landmine detection. Dr. Keller has coauthored over 225 technical publications.

Professor Keller is a Fellow of the Institute of Electrical and Electronics Engineers (IEEE), is a Distinguished Lecturer for the IEEE Computational Intelligence Society (CIS) and for the Association for Computing Machinery (ACM). He is a past President of the North American Fuzzy Information Processing Society (NAFIPS), a past Editor-in-Chief of the *IEEE Transactions on Fuzzy Systems*, and was the general chair for the 2003 IEEE International Conference on Fuzzy Systems. He is currently the Vice President for Publications for CIS.