



THE CHINESE UNIVERSITY OF HONG KONG
Department of Information Engineering

Seminar

Protein Universe and its applications

by

Professor Stephen S.-T. Yau

Professor of Tsinghua University, Beijing, China

Emeritus Distinguished Professor of University of Illinois, U.S.A.

Date : 29 April, 2016 (Friday)

Time : 4:00pm – 5:00pm

**Venue : Room 1009, William M.W. Mong Engineering Building
The Chinese University of Hong Kong**

Abstract

Current methods cannot tell us what the nature of the protein universe is concretely. Here we develop a novel method to realize the nature of protein universe. We show the protein universe can be realized as a protein space in 60-dimensional Euclidean space by using the distribution information of each amino acid within the protein. Every protein is in one-to-one correspondence with a point in protein space, where proteins with similar properties stay close together. Thus the distance between two points in protein space represents the biological distance of the corresponding two proteins. We also propose a natural graphical representation for inferring phylogeny. The representation is natural and unique based on the biological distances of proteins in protein space. This will solve the fundamental question of how proteins are distributed in the protein universe.

As an application, we shall answer the following fundamental question. What kinds of amino acid sequences could possibly be protein sequences? From all existing databases that we can find, known proteins are only a small fraction of all possible combinations of amino acids. Here we show that when the collection of arbitrary amino acid sequences is viewed in an appropriate geometric context, the protein sequences cluster together. This leads to a new computational test, described here, that has proved to be remarkably accurate at determining whether an arbitrary amino acid sequence can be a protein. Even more, if the results of this test indicate that the sequence can be a protein, and it is indeed a protein sequence, then its identity as a protein sequence is uniquely defined.

Biography

Dr. Stephen S.-T. Yau, Professor of Tsinghua University, Emeritus Distinguished Professor of University of Illinois. He is the Managing Editor and founder of the Journal of Algebraic Geometry since 1992, and the Editor-in-Chief and founder of Communications in Information and Systems from 2001 to the present. He was awarded the Sloan Fellowship in 1980, the Guggenheim Fellowship in 2000, the IEEE Fellow in 2003 and the AMS Fellow Award in 2013. His current research interests include bioinformatics, nonlinear filtering, complex algebraic geometry, CR geometry and singularities theory.

**** ALL ARE WELCOME ****