

# CS Seminar

## ***Object-Centric Approaches in Machine Learning***

**Professor Leonidas Guibas**  
**Stanford University**

**Date:**

March 29, 2019

Friday

2:30 pm

**Venue:**

MB103

Main Building

The University of Hong Kong

**Abstract:**

Deep knowledge of the world is necessary if we are to augment real scenes with virtual entities, or to have autonomous and intelligent agents and artifacts that can assist us in everyday activities, or even carry out tasks entirely independently. One way to factorize the complexity of the world is to associate information and knowledge with stable entities, animate or inanimate, such as a person or a vehicle, etc. In this talk I'll survey a number of recent efforts whose aim is to create and annotate reference representations for objects based on 3D models, with the aim of delivering information to new observations, as needed. The information may relate to object geometry, appearance, articulation, materials, physical properties, affordances, or functionality. We acquire such object knowledge in a multitude of ways, both from crowd-sourcing and from establishing direct links between models and signals, such as images, videos, and 3D scans -- and through these to language and text. The purity of the 3D representation allows us to establish robust maps and correspondences for transferring information among the 3D models themselves -- making our current 3D repository, ShapeNet, a true network. Furthermore, the network can act as a regularizer, allowing us to benefit from the "wisdom of the collection" in performing operations on individual data sets or in map inference between them. This effectively enables us to add missing information to signals through "computational imagination", giving us for example the ability to infer what an occluded part of an object in an image may look like, or what other object arrangements may be possible, based on the world-knowledge encoded in 3D shape and scene repositories. One challenge of the 3D world is that 3D data typically come as point clouds or meshes, which do not have the regular grid structure of image or video data. This makes it challenging to apply the highly successful convolutional deep architectures (CNNs) to 3D data, as CNNs heavily depend on neighborhood regularity for weight sharing and other optimizations. The talk will discuss deep architectures capable of directly processing 3D data, such as the PointNet and PointNet++ networks for point clouds, including more recent developments on lifting 2D data to 3D to recover object pose and shape. Finally, we will briefly examine ways to learn object function from observing multiple action sequences involving objects and to synthesize realistic animations of similar actions in new settings, or continuations of an action in its current setting.

**About the Speaker:**

Leonidas Guibas is the Paul Pigott Professor of Computer Science (and by courtesy, Electrical Engineering) at Stanford University, where he heads the Geometric Computation group. Dr. Guibas obtained his Ph.D. from Stanford University under the supervision of Donald Knuth. His main subsequent employers were Xerox PARC, DEC/SRC, MIT, and Stanford. He is a member and past acting director of the Stanford Artificial Intelligence Laboratory and a member of the Computer Graphics Laboratory, the Institute for Computational and Mathematical Engineering (iCME) and the Bio-X program. Dr. Guibas has been elected to the US National Academy of Engineering and the American Academy of Arts and Sciences, and is an ACM Fellow, an IEEE Fellow and winner of the ACM Allen Newell award and the ICCV Helmholtz prize.

**All are welcome!**

**For enquiries, please call 2859 2180 or email**

**[enquiry@cs.hku.hk](mailto:enquiry@cs.hku.hk)**

**Department of Computer Science**

**The University of Hong Kong**

