



HUMANITY
CENTERED
ROBOTICS
INITIATIVE



Jeffrey M. Siskind Purdue University

"Decoding the Brain to Help Build Machines"



Monday, November 9, 2015
1:00-2:00 pm
CIT Building, Room 368

Humans can describe observations and act upon requests. This requires that language be grounded in perception and motor control. I will present several components of my long-term research program to understand the vision-language-motor interface in the human brain and emulate such on computers.

In the first half of the talk, I will present fMRI investigation of the vision-language interface in the human brain. Subjects were presented with stimuli in different modalities---video, text, and speech---while undergoing fMRI. The scan data is analyzed to allow readout of individual constituent concepts and words---names, nouns, verbs, and prepositions---as well as phrases and entire sentences. This can be done across subjects and across modality; we use classifiers trained on scan data for one subject to read out from another subject and use classifiers trained on scan data for one modality, say text, to read out from scans of another modality, say video or speech. Analysis of this indicates that the brain regions involved in processing the different kinds of constituents are largely disjoint but also largely shared across subjects and modality. This points to a compositional mental semantic representation common across subjects and modalities.

In the second half of the talk, I will use this work to motivate the development of two computational systems. First, I will present a system that can use sentential description of human interaction with previously unseen objects in video to automatically find and track those objects. This is done without any annotation of the objects and without any pretrained object detectors. Second, I will present a system that learns the meanings of nouns and prepositions from video and tracks of a mobile robot navigating through its environment paired with sentential descriptions of such activity. Such a learned language model then supports both generation of sentential description of new paths driven in new environments as well as automatic driving of paths to satisfy navigational instructions specified with new sentences in new environments.

Jeffrey M. Siskind received the B.A. degree from the Technion in 1979, the S.M. degree from MIT in 1989, and the Ph.D. degree from MIT in 1992. He did a postdoctoral fellowship at the University of Pennsylvania from 1992 to 1993. He was an assistant professor at the University of Toronto from 1993 to 1995, a senior lecturer at the Technion in 1996, a visiting assistant professor at the University of Vermont from 1996 to 1997, and a research scientist at NEC Research Institute, Inc. from 1997 to 2001. He joined Purdue University in 2002 where he is currently an associate professor. His research interests include computer vision, robotics, artificial intelligence, neuroscience, cognitive science, computational linguistics, child language acquisition, automatic differentiation, and programming languages and compilers.