Natural Language text and speech processing (Computational Linguistics) is just over 50 years old, and is still continuously evolving—not only in its technical subject matter, but in the basic questions being asked and the style and methodology being adopted to answer them. As unification followed finite-state technology in the 1980s, statistical processing followed that in the 1990s, and large-scale processing is increasingly being adopted (especially for commercial NLP) in this decade, a new and quite interesting trend is emerging: a split of the field into three somewhat complementary and rather different directions, each with its own goals, evaluation paradigms, and methodology. The resource creators focus on language and the representations required for language processing; the learning researchers focus on algorithms to effect the transformation of representation required in NLP; and the large-scale hackers produce engines that win the NLP competitions. But where the latter two trends have a fairly well-established methodology for research and papers, the first doesn’t, and consequently suffers in recognition and funding. In the talk, I describe each trend, provide some examples of the first, and conclude with a few general questions, including: Where is the heart of NLP? What is the nature of the theories developed in each stream (if any)? What kind of work should one choose to do if one is a grad student today?

Bio: Eduard Hovy is a member of the Language Technology Institute in the School of Computer Science at Carnegie Mellon University. He holds adjunct professorships at universities in China, Korea, and Canada, and is co-Director of Research for the DHS Center for Command, Control, and Interoperability Data Analytics. He used to direct the Human Language Technology Group at the Information Sciences Institute of the University of Southern California. Dr. Hovy completed a Ph.D. in Computer Science (Artificial Intelligence) at Yale University in 1987. His research addresses areas in Natural Language Processing, including machine reading of text, question answering, information extraction, automated text summarization, the semi-automated construction of large lexicons and ontologies, and machine translation. Dr. Hovy is the author or co-editor of six books and over 300 technical articles and is a popular invited speaker. In 2001 Dr. Hovy served as President of the Association for Computational Linguistics (ACL) and in 2001–03 as President of the International Association of Machine Translation (IAMT). Dr. Hovy regularly co-teaches courses and serves on Advisory Boards for institutes and funding organizations in Germany, Italy, Netherlands, and the USA.

* LTI colloquium: http://www.cs.cmu.edu/~roni/11700/