Announcement of a Defense of a Dissertation

John C. Martin, Candidate for a Doctor of Philosophy
Dr. Michael W. Berry, Major Professor

Friday, May 20, 2016 at 9:00 AM
Claxton 233

"Quantitative Metrics for Comparison of Hyper-dimensional LSA Spaces for Semantic Differences"

ABSTRACT

Latent Semantic Analysis (LSA) is a mathematically based machine learning technology that has demonstrated success in numerous applications in text analytics and natural language processing. The construction of a large hyper-dimensional space, a LSA space, is central to the functioning of this technique, serving to define the relationships between the information items being processed. This hyper-dimensional space serves as a semantic mapping system that represents learned meaning derived from the input content. The meaning represented in an LSA space, and therefore the mappings that are generated and the quality of the results obtained from using the space, is completely dependent on the content used to construct the space. It can be easily observed that modifying the content used to build a LSA space changes the meaning represented by the space, but in current practice the impact of these changes upon the overall body of meaning represented by the space is not understood. The research described here seeks to identify the impact of changes in the content of a LSA space on the meaning represented by that space through the development of quantitative measures. These measures will facilitate the comparison of different LSA spaces to assess their degree of semantic similarity. This insight will in turn provide reasoning leverage for answering questions about the characteristics of LSA spaces related to the overall body of meaning that they represent.