Learning Probabilistic Scripts for Text Understanding
Dr. Ray Mooney

Abstract
In natural language processing, scripts encode knowledge of stereotypical event sequences
countered in text and can be used to improve text comprehension and draw inferences that
are naturally made by human readers. Schank introduced the idea of scripts in the 1970's as a
form of non-probabilistic knowledge; however, Chambers and Jurafsky (2008) and other
recent work has attempted to learn probabilistic script models from large corpora of raw text.
We have recently developed two improved statistical script induction methods, the first
extends previous single verb-argument models to multiple verb arguments; and the second
utilizes Long Short Term Memory (LSTM) recurrent neural networks (which have recently
demonstrated impressive performance on machine translation and image description). We
evaluate these new models using Chambers and Jurafsky's "narrative cloze" task in which an
event (represented as a verb and its arguments) is deleted, and the system must correctly
infer it from the other events in the document. Our new script models demonstrate significantly
improved performance on this evaluation compared to previous methods.

Bio
Raymond J. Mooney is a Professor in the Department of Computer Science at the University
of Texas at Austin. He received his Ph.D. in 1988 from the University of Illinois at
Urbana/Champaign. He is an author of over 150 published research papers, primarily in the
areas of machine learning and natural language processing. He was the President of the
International Machine Learning Society from 2008-2011, program co-chair for AAAI 2006,
general chair for HLT-EMNLP 2005, and co-chair for ICML 1990. He is a Fellow of the
American Association for Artificial Intelligence, the Association for Computing Machinery, and
the Association for Computational Linguistics and the recipient of best paper awards from
AAAI-96, KDD-04, ICML-05 and ACL-07.