

Preface

The majority of human knowledge is encoded in text, and much of this text is available in machine readable form on the web. But to machines, the knowledge encoded in the texts they read remains inaccessible. Significant progress has been made in such basic areas of language processing as morphological analysis, syntactic parsing, proper name recognition, and logical form extraction. This has already advanced information extraction and filtering capabilities, as a variety of current application systems demonstrate. Still, intelligent machines of today cannot yet claim to be able to generate semantic representations on the scale and of the depth sufficient to support automatic reasoning, a situation often blamed on the knowledge acquisition bottleneck.

The goal of this symposium is to stimulate discussion and open exchange of ideas about two aspects of making texts semantically accessible to and processable by machines. The first, learning by reading, relates to automatically extracting machine-understandable (machine-tractable) knowledge from text. The second, learning to read, is related to automating the process of knowledge extraction required to acquire and expand resources (for example, ontologies and lexicons) that facilitate learning by reading. There is a clear symbiotic relationship between these two aspects — expanding knowledge resources enables systems that extract knowledge from text to improve at that task over time and vice versa. Given significant diversity in topics, terminology, and writing styles, learning to read will be crucial to large-scale deployment of systems that learn by reading.

— *Sergei Nirenburg*
(University of Maryland, Baltimore County)

— *Tim Oates*
(University of Maryland, Baltimore County)