Machine-Readable Dictionaries and Natural Language Processing

Orin Hargraves

Independent Lexicographer

When first developed, machine-readable dictionaries were hailed as a great step forward for automated word sense disambiguation. In fact they have proved disappointing to many in NLP and have even been denounced as useless by some. This talk looks at which elements of dictionary data are of use in an NLP system and how dictionary definition structure can abet or hinder tasks in NLP. Several tools available to lexicographers and other researchers will be surveyed with a view to their usefulness in bridging the gap between NLP and dictionary databases more effectively.

Biography

Orin Hargraves is an American lexicographer and writer. He is the Dictionary Society of North America instructor at the 2011 Institute. His language reference works include *Slang Rules!: A Practical Guide for English Learners* (Merriam-Webster, 2008), and *Mighty Fine Words and Smashing Expressions: Making Sense of Transatlantic English* (Oxford University Press, 2003). He contributes regularly to dictionaries and other language reference works published by Berlitz, Cambridge University Press, HarperCollins, Longman, Chambers Harrap, Langenscheidt, Merriam-Webster, and Oxford University Press, among others. Most recently he was chief lexicographer for the *Scholastic Children's Dictionary* (2010). He is the American English consultant for the forthcoming third edition of *Collins CoBuild*

English Grammar (2011).