Justin Betteridge  PhD student

Never-Ending Language Learning,  Information Extraction

Advisor: Tom Mitchel
Research project: Read the Web
rtw.ml.cmu.edu

Read the Web
Research Project at Carnegie Mellon University

NELL: Never-Ending Language Learning

Can computers learn to read? We think so. "Read the Web" is a research project that attempts to create a computer system that learns over time to read the web. Since January 2010, our computer system called NELL (NeverEnding Language Learner) has been running continuously, attempting to perform two tasks each day:

- First, it attempts to "read," or extract facts from text found in hundreds of millions of web pages (e.g., playsInstrument(George Harrison, guitar)).
- Second, it attempts to improve its reading competence, so that tomorrow it can extract more facts from the web, more accurately.

At present, NELL has accumulated a knowledge base of 497,773 beliefs that it has read from various web pages. It is not perfect, but NELL is learning. You can track NELL’s progress below or @cmuNELL on Twitter, browse and download its knowledge base, read more about our technical approach, or join the discussion group.