



Distinguished Lecture Series

Never-Ending Learning



Wednesday, November 17th, 2010 10:00am
Auditorium 106 at New IIS Building

Tom M. Mitchell

E. Fredkin University Professor
Chair, Machine Learning Department, School of
Computer Science, Carnegie Mellon University

Abstract

What would it take to develop machine learners that run forever, each day improving their performance and also the accuracy with which they learn? This talk will describe our attempt to build a never-ending language learner, NELL, that runs 24 hours per day, forever, and that each day has two goals: (1) extract more structured information from the web to populate its growing knowledge base, and (2) learn to read better than yesterday, by using previously acquired knowledge to better constrain its subsequent learning.

The approach implemented by NELL is based on two key ideas: coupling the semi-supervised training of hundreds of different functions that extract different types of information from different web sources, and automatically discovering new constraints that more tightly couple the training of these functions over time. NELL has been running nonstop since January 2010 (follow it at <http://rtw.ml.cmu.edu>), and has extracted a knowledge base containing approximately 440,000 beliefs as of October 2010. This talk will describe NELL, its successes and its failures, and use it as a case study to explore the question of how to design never-ending learners.

For more information: <http://www.iis.sinica.edu.tw/>

