Lifelong Credit Assignment with the Success-Story Algorithm

Jürgen Schmidhuber
IDSIA, Galleria 2, 6928 Manno-Lugano, Switzerland
University of Lugano & SUPSI, Switzerland

Abstract

Consider an embedded agent with a self-modifying, Turing-equivalent policy that can change only through active self-modifications. How can we make sure that it learns to continually accelerate reward intake? Throughout its life the agent remains ready to undo any self-modification generated during any earlier point of its life, provided the reward per time since then has not increased, thus enforcing a lifelong success-story of self-modifications, each followed by long-term reward acceleration up to the present time. The stack-based method for enforcing this is called the success-story algorithm. It fully takes into account that early self-modifications set the stage for later ones (learning a learning algorithm), and automatically learns to extend self-evaluations until the collected reward statistics are reliable... a very simple but general method waiting to be re-discovered! Time permitting, I will also briefly discuss more recent mathematically optimal universal maximizers of lifelong reward, in particular, the fully self-referential Gödel machine.

Note: This is a summary of earlier work (Schmidhuber 1994; Schmidhuber, Zhao, and Schraudolph 1997; Schmidhuber 1987; 1993; Wiering and Schmidhuber 1996; Schmidhuber, Zhao, and Wiering 1997; Schmidhuber and Zhao 1997; Hutter 2005; Schmidhuber 2006; 2005; 2009).

References


