

# Tom Mitchell - Never-ending learning (2018)

Tom Rochette <tom.rochette@coreteks.org>

March 22, 2020 — [57bd6871](#)

## 0.1 Context

## 0.2 Learned in this study

## 0.3 Things to explore

- Given that the problem that NELL seems to be tackling has some resemblance to playing Go, is it possible to employ the same techniques used by AlphaGo/Zero?
  - Is there a way to convert their problem into a problem that would benefit from doing some kind of
- Devise a metric which allows the agent to focus on where it needs the most improvement (deliberate practice/training)
- A key open research question is how the learning agent might itself evolve a useful curriculum of learning tasks

## 1 Overview

## 2 Notes

### 2.1 Abstract

- We suggest that people learn better than computers precisely because (they learn many different types of knowledge from diverse experiences over many years and become better learners overtime whereas machine learning systems learn just a single function or data model)
- We suggest a key direction for machine learning research is to develop software architectures that enable intelligent agents to also learn many types of knowledge, continuously over many years, and to become better learners over time

### 2.2 1. Introduction

- Humans learn cumulatively: as they learn one thing this new knowledge helps them to more effectively learn the next, and if they revise their beliefs about the first then this change refines the second
- The thesis of our research is that we will never truly understand machine or human learning until we can build computer programs that, like people
  - learn many different types of knowledge of functions
  - from years of diverse, mostly self-supervised experience
  - in a staged curricular fashion, where previously learned knowledge enables learning further types of knowledge,
  - where self-reflection and the ability to formulate new representations and new learning tasks enable the learner to avoid stagnation and performance plateaus

## 2.3 3. Never-Ending Learning

- We define a never-ending learning agent to be a system that, like humans, learns many types of knowledge, from years of diverse and primarily self-supervised experience, using previously learned knowledge to improve subsequent learning, with sufficient self-reflection to avoid plateaus in performance as it learns
- The never-ending learning problem faced by the agent consists of a collection of learning tasks, and constraints that couple their solutions
- We define a never-ending learning problem  $\mathcal{L}$  to be an ordered pair consisting of
  - a set  $L = \{L_i\}$  of learning tasks, where the  $i$ th learning task  $L_i = \langle T_i, P_i, E_i \rangle$  is to improve the agent’s performance, as measured by performance metric  $P_i$ , on a given performance task  $T_i$ , through a given type of experience  $E_i$
  - a set of coupling constraints  $C = \{\langle \phi_k, V_k \rangle\}$  among the solutions to these learning tasks, where  $\phi_k$  is a real-valued function over two or more learning tasks, specifying the degree of satisfaction of the constraint, and  $V_k$  is a vector of indices over learning tasks, specifying the arguments to  $\phi_k$
- Each performance task  $T_i$  is a pair  $T_i = \langle X_i, Y_i \rangle$  defining the domain and range of a function to be learned  $f_i^* : X_i \rightarrow Y_i$
- The performance metric  $P_i : f \rightarrow \mathbb{R}$  defines the optimal learned function  $f_i^*$  for the  $i$ th learning task:

$$f_i^* \equiv \arg \max_{f \in F_i} P_i(f)$$

where  $F_i$  is the set of all possible functions from  $X_i$  to  $Y_i$

## 2.4 6. NELL’s Learning Methods and Architecture

### 2.4.1 6.4 Self-Reflection and Self-Evaluation

- Under what conditions can unlabeled data be used to estimate accuracy of learned functions?
  - We have found that there are conditions under which the observed consistency among different learned functions applied to unlabeled data can be used to derive high precise estimates of accuracies of these functions, and that these methods work well for accuracy estimation in NELL

## 2.5 8. Discussion

- Four useful design features that have led to the successes it has had - design features we recommend for any never-ending system:
  - To achieve successful semi-supervised learning, couple the training of many different learning tasks
  - Allow the agent to learn additional coupling constraints
  - Learn new representations that cover relevant phenomena beyond the initial representation
  - Organize the set of learning tasks into an easy-to-increasingly-difficult curriculum
- Additional areas for research into never-ending learning agents:
  - Self reflection and an explicit agenda of learning subgoals
  - Pervasive plasticity
  - Representation and learning

## 3 See also

## 4 References

- Mitchell, Tom, et al. “Never-ending learning.” Communications of the ACM 61.5 (2018): 103-115.