Artificial Intelligence 219 (2015) 67–91

Contents lists available at ScienceDirect

Artificial Intelligence

www.elsevier.com/locate/artint

Integrating representation learning and skill learning in a human-like intelligent agent

Nan Li *, Noboru Matsuda, William W. Cohen, Kenneth R. Koedinger

5000 Forbes Ave, Pittsburgh, PA 15232, USA

A R T I C L E   I N F O

Article history:
Received 18 April 2012
Received in revised form 2 July 2014
Accepted 5 November 2014
Available online 4 December 2014

Keywords:
Agent learning
Representation learning
Student modeling

A B S T R A C T

Building an intelligent agent that simulates human learning of math and science could potentially benefit both cognitive science, by contributing to the understanding of human learning, and artificial intelligence, by advancing the goal of creating human-level intelligence. However, constructing such a learning agent currently requires manual encoding of prior domain knowledge; in addition to being a poor model of human acquisition of prior knowledge, manual knowledge-encoding is both time-consuming and error-prone. Previous research has shown that one of the key factors that differentiates experts and novices is their different representations of knowledge. Experts view the world in terms of deep functional features, while novices view it in terms of shallow perceptual features. Moreover, since the performance of learning algorithms is sensitive to representation, the deep features are also important in achieving effective machine learning. In this paper, we present an efficient algorithm that acquires representation knowledge in the form of "deep features", and demonstrate its effectiveness in the domain of algebra as well as synthetic domains. We integrate this algorithm into a machine-learning agent, SimStudent, which learns procedural knowledge by observing a tutor solve sample problems, and by getting feedback while actively solving problems on its own. We show that learning "deep features" reduces the requirements for knowledge engineering. Moreover, we propose an approach that automatically discovers student models using the extended SimStudent. By fitting the discovered model to real student learning curve data, we show that it is a better student model than human-generated models, and demonstrate how the discovered model may be used to improve a tutoring system's instructional strategy.

1. Introduction

One of the fundamental goals of artificial intelligence is to understand and develop intelligent agents that simulate human-like intelligence. A considerable amount of effort [1–3] has been put toward this challenging task. Further, education in the 21st century will be increasingly about helping students not just learn content but to become better learners. Thus, we have a second goal of improving our understanding of how humans acquire knowledge and how students vary in their abilities to learn.

* Corresponding author.
E-mail addresses: nli1@cs.cmu.edu (N. Li), Noboru.Matsuda@cs.cmu.edu (N. Matsuda), wcohen@cs.cmu.edu (W.W. Cohen), koedinger@cmu.edu (K.R. Koedinger).

http://dx.doi.org/10.1016/j.artint.2014.11.002

To contribute to both goals, there have been recent efforts [4–7] in developing intelligent agents that model human learning of math, science, or a second language. Although such agents produce intelligent behavior with less human knowledge engineering than before, there remains a non-trivial element of knowledge engineering in the encoding of the prior domain knowledge given to the simulated student agent at the start of the learning process. For example, to build an algebra learning agent, the agent developer needs to provide prior knowledge by coding functions that describe, for instance, how to extract a coefficient or how to add two algebraic terms. Such manual encoding of prior knowledge can be time-consuming and the constructed prior knowledge may not naturally correspond with a human student's prior knowledge.

Since real students entering a course do not usually have substantial domain-specific or domain-relevant prior knowledge, it is not realistic in a model of human learning to assume this knowledge is given rather than learned. For example, for students learning about algebra, we cannot assume that they all know beforehand what a coefficient is, or what the difference between a variable term and a constant term is. An intelligent system that models automatic knowledge acquisition with a small amount of prior knowledge could be helpful both in reducing the effort in knowledge engineering intelligent systems and in advancing the cognitive science of human learning.

Previous work in cognitive science [8,9] showed that one of the key factors that differentiates experts and novices in a field is their different prior knowledge of world state representation. Experts view the world in terms of deep functional features (e.g., coefficient and constant in algebra), while novices only view in terms of shallow perceptual features (e.g., integer in an expression). Deep features are often domain-specific, whereas shallow perceptual features are domain-independent. Having the correct representation of the deep features aids the process of solving the domain task. For example, in algebra, students need to learn to encode equation input into "terms" and "coefficients". A shallow feature encoding of a coefficient (e.g., of the "*5*" in "*5x*") is as a number before a letter. A deep feature encoding requires the learner to develop knowledge, which may include implicit perceptual processing capabilities, to recognize coefficients more generally such as the "*-5*" in "*-5x*", the "*a*" in "*ax*", the "*3*" in "*3(x+2)*", the "*-1*" in "*-x*". In general, experts develop deep feature knowledge that allows them to see the world in the way novices do not – expert readers see "run" as a word whereas novices see letters or just lines, experts in physics see force contact points whereas novices see blocks and inclined planes, chess experts see configurations of pieces like a knight fork whereas novices see pieces. Such deep feature perception knowledge is learned for specific domains, perhaps as much by implicit experience as by explicit instruction. In algebra, students learn to see terms and coefficients in equations building upon more general prior knowledge of numbers. That prior knowledge may be the basis for initial shallow feature encoding as in the example above. In general, we consider deep features as part of representation knowledge. Representation knowledge organizes low-level perceptual input into a structured form that assists the agent to understand and solve problems in a particular domain. Even if the same perceptual input was given, for different problem solving tasks in different domains, the ideal representation of the world can be different for different tasks. Deep features can be viewed as the key features that differentiate a well-structured representation from a poorly-structured representation.

Deep feature learning is a major component of human expertise acquisition, but has not received much attention in AI. Learning deep features changes the representation on which future learning is based and, by doing so, improves future learning. However, how these deep features are acquired is not clear. Therefore, we have recently developed a learning algorithm that acquires deep features automatically with only domain-independent knowledge (e.g., what is an integer) as input [10]. We evaluated the effectiveness of the algorithm in learning deep features, but not its impact on future skill learner.

In order to evaluate how the deep feature learner could affect future learning of an intelligent agent, in this paper, we integrated this deep feature learning algorithm into *SimStudent* [11], an agent that learns problem-solving skills by example and by feedback on performance. The original SimStudent relies on a hand-engineered representation that encodes an expert representation given as prior knowledge. This limits its ability to model novice students. The extended SimStudent first acquires the representation of the problems using the deep feature learner. Then, it makes use of the learned representation to acquire skill knowledge in later tasks. Integrating the deep feature learner into the original SimStudent both reduces the amount of engineering effort and builds a better model of student learning.

We show that the extended SimStudent with better representation learning performs much better than the original SimStudent when neither of them are given domain-specific knowledge. Furthermore, we also show that even compared to the original SimStudent with the domain-specific knowledge, the extended SimStudent is able to learn nearly as well without being given domain-specific knowledge. For the sake of simplicity, we only report experiment results in the algebra domain in this paper, but similar results are also observed in other domains [12]. In addition, we use the extended SimStudent to automatically discover models of real students, and show that the discovered models are better student models than human-generated models [13]. Although not reported here, we further use the extended SimStudent to better understand how problem orders affect learning effectiveness by inspecting SimStudent's learning processes and learning outcomes, which are not easily obtainable from human subjects [14].

To summarize, the main contributions of this paper are two-fold. By integrating representation learning into skill learning, 1) we reduce the amount of knowledge engineering effort required in constructing an intelligent agent; 2) we get a better model of human behavior.

In the following sections, we start with a brief review of SimStudent. We then present the deep feature learning algorithm together with its evaluation results. Next, we describe how to integrate the deep feature learner into SimStudent, and illustrate the algorithm with an example in algebra. After that, we present experimental results for both the original Sim-