Available online at www.sciencedirect.com
ScienceDirect
Neural Networks 20 (2007) 34–47
Neural Networks
www.elsevier.com/locate/neunet

2007 Special Issue

The interaction of implicit learning, explicit hypothesis testing learning and implicit-to-explicit knowledge extraction

Ron Sun[a,*], Xi Zhang[b], Paul Slusarz[b], Robert Mathews[c]

[a] Department of Cognitive Science, Rensselaer Polytechnic Institute, Troy, NY 12180, USA
[b] Department of Computer science, University of Missouri-Columbia, Columbia, MO 65211, USA
[c] Psychology Department, Louisiana State University, Baton Rouge, LA 70803-5501, USA

Received 26 July 2005; accepted 7 July 2006

Abstract

To further explore the interaction between the implicit and explicit learning processes in skill acquisition (which have been tackled before, e.g. in [Sun, R., Merrill, E., & Peterson, T. (2001). From implicit skill to explicit knowledge: A bottom-up model of skill learning. *Cognitive Science*, *25(2)*, 203–244; Sun, R., Slusarz, P., & Terry, C. (2005). The interaction of the explicit and the implicit in skill learning: A dual-process approach. *Psychological Review*, *112(1)*, 159–192]), this paper explores details of the interaction of different learning modes: implicit learning, explicit hypothesis testing learning, and implicit-to-explicit knowledge extraction. Contrary to the common tendency in the literature to study each type of learning in isolation, this paper highlights the interaction among them and various effects of the interaction on learning, including the synergy effect. This work advocates an integrated model of skill learning that takes into account both implicit and explicit learning processes; moreover, it also uniquely embodies a bottom-up (implicit-to-explicit) learning approach in addition to other types of learning. The paper shows that this model accounts for various effects in the human behavioural data from the psychological experiments with the process control task, in addition to accounting for other data in other psychological experiments (which has been reported elsewhere). The paper shows that to account for these effects, implicit learning, bottom-up implicit-to-explicit extraction and explicit hypothesis testing learning are all needed.

Keywords: Cognitive modelling; Cognitive science; Psychology; Skill learning; Implicit learning; Neural networks; Backpropagation; Reinforcement learning

1. Introduction

The role of implicit learning in skill acquisition and the distinction between implicit and explicit learning have been widely recognized in recent years (see, e.g. Proctor and Dutta (1995), Reber (1989), Stanley, Mathews, Buss, and Kotler-Cope (1989), Willingham, Nissen, and Bullemer (1989)). Although implicit learning as well as explicit learning have been actively investigated, the complex interaction between the implicit and the explicit and the importance of this interaction have not been widely recognized; such interaction has traditionally been downplayed or ignored, with only a few notable exceptions (e.g. Mathews et al. (1989), Sun, Merrill, and Peterson (1998, 2001), Sun (1999), Sun, Slusarz, and Terry (2005)). Research has been focused on showing the *lack* of explicit learning in various learning settings (see especially Lewicki, Czyzewska, and Hoffman (1987)) and on the controversies stemming from such claims. Similar oversight is also evident in computational simulation models of implicit learning (with few exceptions such as Cleeremans (1994), Sun et al. (2001)).

Despite the lack of studies of interaction, it has been gaining recognition that it is difficult to find a situation in which only one type of learning is engaged (Reber (1989), Seger (1994), but see Lewicki et al. (1987)). Various indications of the interaction are scattered throughout the literature. For instance, Stanley et al. (1989) found that under some circumstances concurrent verbalization (which leads to generating more explicit knowledge) could help to improve human subjects' (mostly implicit) performance in a process control task (the detail of which will be explained later). Ahlum-Heath and DiVesta (1986) also found that verbalization led to better performance in learning Tower of Hanoi. (However, note that, as Reber (1976), Reber (1989), Sun et al. (2001) pointed

* Corresponding author. Tel.: +1 518 276 3409.
E-mail address: rsun@rpi.edu (R. Sun).

0893-6080/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neunet.2006.07.002

out, verbalization and the resulting explicit knowledge might also hamper implicit learning, especially when too much verbalization induced an overly explicit learning mode in human subjects performing a task that was not suitable for learning in an explicit way.)

As variously demonstrated by Berry and Broadbent (1988), Stanley et al. (1989), Reber, Kassin, Lewis, and Cantor (1980), verbal instruction given prior to learning can also facilitate or hamper task performance. One type of instruction was to encourage explicit search by human subjects for regularities that might aid in task performance. For example, Reberet al. (1980) found that, depending on the ways stimuli were presented, explicit search might help or hamper performance. Owen and Sweller (1985), Schooler, Ohlsson, and Brooks (1993) found that explicit search hindered learning. Another type of instruction was explicit how-to instruction that told human subjects specifically how the tasks should be performed, including providing detailed information concerning regularities in stimuli. Stanley et al. (1989) found that such instructions helped to improve performance significantly.

In terms of the relation between implicit and explicit knowledge acquired during learning, there is some evidence that implicit and explicit knowledge may develop independently under some circumstances. Willingham et al. (1989), for example, reported some psychological data that were consistent with the parallel development of implicit and explicit knowledge. By using two different measures (with varying criterion levels) for assessing the two types of knowledge respectively, they compared the time course of implicit and explicit learning. It was shown that implicit knowledge might be acquired in the absence of explicit knowledge and vice versa.

There are also cases where a subject's performance improves earlier than explicit knowledge that can be verbalized by the subject (Stanley et al., 1989). For instance, in process control tasks (to be detailed later), while the performance of the subjects quickly rose to a high level, their verbal knowledge improved far more slowly: The subjects could not provide usable verbal knowledge (for novice subjects to use) until near the end of their training (see Stanley et al. (1989)). It appears that in these tasks, it is much easier to acquire implicit skills than to acquire explicit knowledge, and hence there is a delay in the development of explicit knowledge. Bowers, Regehr, Balthazard, and Parker (1990) also showed delayed learning of explicit knowledge. When subjects were given certain partial patterns to complete, they first showed implicit recognition of proper completion (though they did not have explicit recognition). Their implicit recognition improved over time until eventually an explicit recognition was achieved. This phenomenon was also demonstrated by Reber and Lewis (1977) in artificial grammar learning. In all of these cases, as suggested by Stanley et al. (1989), Seger (1994), and Sun et al. (2001), due to the fact that explicit knowledge lags behind but improves along with implicit knowledge, explicit knowledge is in a way 'extracted' from implicit knowledge. Learning of explicit knowledge appears to occur through the (delayed) *explication* of implicit knowledge.

In the development of cognitive architectures, the distinction between procedural and declarative knowledge has been proposed for a long time, and adopted by many in the field, (e.g. Anderson (1993)), although not all (e.g. Rosenbloom, Laird, and Newell (1993)). The distinction maps roughly onto the distinction between the explicit and implicit knowledge, because procedural knowledge is generally inaccessible while declarative knowledge is generally accessible and thus explicit (but also see differing views in, e.g. Anderson and Lebiere (1998), Lebiere, Wallach, and Taatgen (1998)). However, in work on cognitive architectures, the focus has been almost exclusively on *top-down* models (that is, learning first explicit knowledge and then implicit knowledge on the basis of the explicit knowledge), the *bottom-up* direction (that is, learning first implicit knowledge and then explicit knowledge, or learning both in parallel) has been largely ignored. As pointed out earlier, there is work that did demonstrate the parallel development of the two types of knowledge or the extraction of explicit knowledge from implicit knowledge (e.g. Rabinowitz and Goldberg (1995), Stanley et al. (1989), Willingham et al. (1989)), contrary to the common top-down approaches in developing cognitive architectures.

With regard to the interaction between implicit and explicit processes, many issues arise: (1) How can we best model implicit and explicit learning processes computationally? (2) How is bottom-up learning (implicit-to-explicit transition) possible and how can it be realized computationally (e.g. Stanley et al. (1989), Sun et al. (2001))? (3) How do these different types of learning (explicit, implicit, and implicit-to-explicit) interact and contribute to overall skill learning? (4) How do different types of knowledge (acquired from these different types of learning) interact during skilled performance and what is the impact of that interaction on performance? For example, the synergy of the two may be produced, as shown by Sun et al. (2001).

In order to further understand the interaction between the implicit and explicit learning processes in skill acquisition (Sun et al., 2001; Sun & Zhang, 2002, 2003), this paper explores finer details of the interaction of different learning modes: that is, implicit learning, learning through explicit hypothesis testing, and implicit-to-explicit knowledge extraction (bottom-up learning). Contrary to the prevailing tendency in the literature to study each type of learning in isolation, this paper highlights the interaction among them and various effects of the interaction on learning, including the synergy effect. This work advocates an integrated model of skill learning that takes into account both implicit and explicit learning processes. Moreover, it embodies both a bottom-up learning approach (first learning implicit knowledge and then explicit knowledge on the basis of implicit knowledge) and an explicit hypothesis testing learning approach towards learning explicit knowledge. The paper shows that this model accounts for various effects of the implicit/explicit interaction demonstrated in the psychological literature.

In this work, we choose to use the human psychological data of the process control task from Stanley et al. (1989). The data are typical of human performance in process control