

# Steps towards a Challenging Teachable Agent

Annika Silvervarg<sup>1</sup>, Camilla Kirkegaard<sup>1</sup>, Jens Nirme<sup>2</sup>, Magnus Haake<sup>2</sup>,  
and Agneta Gulz<sup>1</sup>

<sup>1</sup> Dept. Computer and Information Science, Linköping University, Linköping, Sweden  
{annika.silvervarg,camilla.kirkegaard,agneta.gulz}@liu.se

<sup>2</sup> Cognitive Science, Lund University, Kungshuset, Lundagård, Lund, Sweden  
{jens.nirme,magnus.haake}@lucs.lu.se

**Abstract.** This paper presents the first steps towards a new type of pedagogical agent – a Challenger Teachable Agent, CTA. The overall aim of introducing a CTA is to increase engagement and motivation and challenge students into deeper learning and metacognitive reasoning. The paper discusses desired design features of such an agent on the basis of related work and results from a study where 11-year old students interacted with a first version of a CTA in the framework of an educational software for history. The focus is on how students respond when the CTA disagrees and questions their suggestions, and how groups of students, differing in response behavior and in self-efficacy, experience the CTA.

**Keywords:** Teachable agent, challenge, interaction, learning, experience.

## 1 Introduction and Background

This paper explores students' interaction with and responses to a new type of pedagogical agent – a challenger teachable agent (CTA). In brief, a teachable agent (TA) is an embodied computer agent which is taught or trained by a student where AI techniques guide the agent's behaviour based on what the agent is taught [1]. Importantly, a TA has no knowledge to begin with, but the knowledge that it gains reflects, more or less, what it is being taught by the student.

Overall, teachable agents have proven pedagogically powerful as an implementation of the *learning by teaching* pedagogy [2, 3]. Our goal is to boost this pedagogical power even further by introducing a teachable agent with a more explicit agency or "will of its own". To our knowledge, the TAs developed so far do not show much of a "will". They do not, for example, argue with a student on whether a piece of information is adequate for a task or not, or indicate that they find a particular topic uninteresting.

Regarding pedagogical agents in general, Frasson and Aïmeur [4] proposed *troublemaker agents* as a subset of learning companions that would question and challenge a student. Such an agent suggests a solution and then asks the student if she agrees or not. If the student does not agree, the troublemaker will argue for her or his own solution – whether it is correct or not – until the student either agrees or the troublemaker runs out of arguments. The students are thereby encouraged to question

their own knowledge and be more motivated as teachers. Several studies show learning gains from troublemaker agents, particularly for high-achieving students, e.g. [4, 5].

Even one of the seminal papers on teachable agents by Brophy et al. [1] proposed a teachable agent that “may be impetuous, not listen or collaborate well”. The implementation and study presented in this article is, however, the first practical attempt in this direction. In an educational software for learning history, we introduce a CTA that, during learning activities where the CTA and student work together and take turns, questions and challenges the student in various ways.

The main motive for introducing a CTA is to stimulate deep learning (c.f. [5]). By being questioned and challenged at times by his or her TA we hope the student becomes encouraged to think once more and perhaps reorganize or rephrase the material he/she is teaching [6]. We also wish to stimulate metacognitive abilities, i.e. reflection on problem solving and learning, abilities that have a transitive value for students when faced with future challenges [7]. In addition, our previous studies have shown that students frequently ask for a TA “with more of an attitude”. It can be boring to interact with an agent that is always positive, compliant and cheerful – and such agents are weak in believability [8].

In the longer run, we are implementing the following set of challenging behaviours in our CTA.

1. The CTA inducing confusion or cognitive disequilibrium by contradicting the student with the aim to provoke the student to reflect on what is true, thereby processing the material at a deeper level [9].
2. The CTA requesting clarification of a solution in a learning activity, thus creating opportunities for debating the study material before accepting it, potentially prompting a desire in the student to share meaning and be understood [10]. This kind of teachable agent behaviour is represented in *SimStudent* [11] where the TA interacts with the student in natural language while solving equations. The TA tries to solve a problem step-by-step and the user verifies the correctness of each step. The TA can ask follow-up questions during this process, which forces the student to reflect on the concepts in the current problem and to show how well he/she understands the material.
3. The CTA occasionally introducing erroneous facts during the learning activities. This will hopefully provoke the student to justify his/her answers [4]. Training to distinguish between right and wrong solutions is also a means to achieve higher confidence in a domain.
4. The CTA prompting the student to choose a task at a more challenging level.

For the study presented in this paper we focus on this first aspect, i.e. the CTAs questioning of the students’ proposals – both in cases where the students’ proposal is correct and when it is not. To question someone you are collaborating with means to introduce a conflict. Several authors discuss the educational potential of conflict or dissonance in other areas than that of teachable agents. For instance, one of the five modes that Weinberger and Fischer [12] lift forth in their analysis of different “social modes of co-construction” in students’ collaborative activities is that of