Modelling human teaching tactics and strategies for tutoring systems

Article (Unspecified)

MODELLING HUMAN TEACHING TACTICS AND STRATEGIES FOR TUTORING SYSTEMS

Benedict du Boulay and Rosemary Luckin, *Human Centred Technology Research Group, School of Cognitive and Computing Sciences, University of Sussex, BN1 9QH, UK.*
bend@cogs.susx.ac.uk, rosel@cogs.susx.ac.uk

Abstract. One of the promises of ITSs and ILEs is that they will teach and assist learning in an intelligent manner. Historically this has tended to mean concentrating on the interface, on the representation of the domain and on the representation of the student's knowledge. So systems have attempted to provide students with reifications both of what is to be learned and of the learning process, as well as optimally sequencing and adjusting activities, problems and feedback to best help them learn that domain.

We now have embodied (and disembodied) teaching agents and computer-based peers, and the field demonstrates a much greater interest in metacognition and in collaborative activities and tools to support that collaboration. Nevertheless the issue of the teaching competence of ITSs and ILEs is still important, as well as the more specific question as to whether systems can and should mimic human teachers. Indeed increasing interest in embodied agents has thrown the spotlight back on how such agents should behave with respect to learners.

In the mid 1980s Ohlsson and others offered critiques of ITSs and ILEs in terms of the limited range and adaptability of their teaching actions as compared to the wealth of tactics and strategies employed by human expert teachers. So are we in any better position in modelling teaching than we were in the 80s? Are these criticisms still as valid today as they were then?

This paper reviews progress in understanding certain aspects of human expert teaching and in developing tutoring systems that implement those human teaching strategies and tactics. It concentrates particularly on how systems have dealt with student answers and how they have dealt with motivational issues, referring particularly to work carried out at Sussex: for example, on responding effectively to the student's motivational state, on contingent and Vygotskian inspired teaching strategies and on the plausibility problem. This latter is concerned with whether tactics that are effectively applied by human teachers can be as effective when embodied in machine teachers.

INTRODUCTION

Intelligent Teaching Systems and Intelligent Learning Environments initially evolved in a rather lop-sided way. First, much effort was put into developing highly detailed models of particular domains. So, for example, SOPHIE in its various versions embodied a highly detailed representation of an electronic circuit at various levels of granularity, from a single device such as a resistor, via a functional sub-structure up to a complete power supply (Brown and Burton, 1975). It also modelled diagnostic tactics and strategy and could react sensibly when students exhibited less than optimal trouble-shooting behaviour. It also had, for its time, good natural language capabilities and could respond to a wide range of domain specific questions and commands (Burton and Brown, 1977). Other systems were able to exploit their domain knowledge, including knowledge of misunderstandings of the domain in order to make fine-grained diagnostic judgements about students. For example, Debuggy (Burton, 1982) and more recent systems like it were capable of building a highly detailed (student) model of an individual's subtraction behaviour, but left it to human teachers to embark on appropriate remediation. Indeed whether there was any value in undertaking such fine-grained diagnosis was itself questioned, as reteaching the whole procedure rather than just the incorrectly understood portion seems just as effective (Sleeman et al., 1989).

Second, teaching involves a wide variety of communicative activities such as explaining, persuading, arguing, demonstrating, describing and so on, and these are skills that are also used in other than educational contexts. One could imagine a teaching system that implemented this more general communicative competence and then specialised it as needed for the particular educational context at the time. The theories of teaching that were implemented in machine teachers[1] were not grounded in such general communicative competence (because it was beyond the state of the art) but necessarily treated "teaching" as an isolated and largely self-contained skill.

An example of one of these isolated and self-contained skills was Socratic Tutoring, a method of teaching based on asking the student a series of carefully constructed questions that would lead students to recognize and fix gaps and inconsistencies in what they know of a domain (Collins et al., 1975). Another example is provided by the various systems produced by Anderson and his colleagues which monitored students' problem-solving in a fine-grained manner and had the capability of reacting immediately if the student departed from the path that an ideal student would have followed (see e.g., Anderson and Reiser, 1985). This endowed machine teachers of that era with a certain communicative brittleness that could undermine their other skills.

Of course, there were attempts to build tutorial systems with more versatile educational communicative competence. One such system was GUIDON (Clancey, 1982) which incorporated rules for "selecting discourse patterns", for "choosing domain knowledge" and for "maintaining the communication model". Within the category of discourse pattern there were rules for responding to a student hypothesis, which resembled Socratic Tutoring, as well as rules for dealing with other aspects of the interaction.

Another such system was the Meno-tutor which incorporated a Discourse Management Network (Woolf, 1988). Indeed, there are various similarities between this network and Ohlsson's taxonomy (see Figure 1). This system could make use of the current discourse context to distinguish and execute a range of different kinds of tutorial tactics (for example, briefly acknowledging a student's incorrect answer) and strategic rules (for example, undertaking a series of shallow questions about a variety of topics).

So overall there was rather uneven progress in the following areas, with most systems having rather a restricted repertoire of teaching actions, and work concentrating on (ii) below:

i. The development of a varied repertoire of teaching actions.

ii. The development of effective strategic and tactical means-ends rules for the deployment of the teaching actions.

iii. The development of such basic, communicative skills and competence as explaining arguing, convincing, cajoling, detecting misunderstandings, dealing with interruptions and side issues etc.

iv. The development of theories of motivation and affect that would enable the judicious change of topic, use of a joke, imposition of a threat, offer of praise and so on.

The kind of criticism that was levelled at machine teachers was the same as that often levelled at AI in general, namely that they tended to concentrate on toy worlds (albeit often highly detailed toy worlds) and that they tended to degrade badly when moved outside their own sphere

[1] The term "machine teachers" is used as a general description of systems that adjust themselves to the needs of their students. This may mean posing a problem, evaluating an answer or adjusting the level of help, but it could also includes adjusting some aspect of (say) a simulation in an intelligent learning environment, to increase the likelihood that productive explorations are undertaken.

of competence (see e.g., Dreyfus, 1979). This meant that the teaching style of most machine teachers was geared towards a rather “convergent”, “syllabus bound” teaching and diagnostic style (see e.g., Ohlsson, 1987). By contrast, a human teacher is able to integrate topics across widely differing domains, change style and approach as the occasion demands, appeal to common-sense knowledge and reasoning and use all the communication and social skills at his or her disposal. A machine teacher often appears plodding and relentless, dominated by its own domain-specific knowledge and unable to deploy any of that change of pace and perspective that makes good teaching what it is.

Another criticism of machine teachers is that they tended to embody a model of teaching in which the teacher knows best. It was not that the systems of the time were wholly concerned with “transmitting” knowledge and maintaining agency in interactions with their users: some excellent systems were arranged as learning environments which reacted intelligently to moves instigated by users. But they were not the kinds of systems that carried out the more facilitating roles of teachers, such as helping students to work more effectively together, helping students reflect on what they had learned and done or guiding students in open-ended project work. There were good reasons for this, namely that modelling such facilitating skills needed artificial intelligence abilities that were (and largely still are) beyond the state of the art. So it’s not that the designers of such systems had an impoverished view of education; it was much more a case of doing what was possible.

This paper is divided into four more sections after this Introduction. The next section (2) expands the critique above by giving a brief account of Ohlsson’s and others’ analyses of the restricted teaching capability and versatility of systems built up to about the mid-eighties. It cites further examples from that period. Section 3 then examines three methodologies for developing teaching capability and versatility. The most important of these methodologies is the observation of human teachers, and examples of such work are described, concentrating in particular on how teachers deal with student errors and how teachers motivate students. Teaching in all its varied forms covers many more factors than just these two, but these have been chosen as representative of the modelling effort of artificial intelligence in education.

In order to see what progress has been made since the mid-eighties, Section 4 draws the threads together from the historical analysis and examines a number of contemporary systems that attempt to embody clever teaching tactics or strategies, including one that attempts to deal with motivational issues. The concluding section offers some thoughts on the degree of progress made.

We must offer two immediate disclaimers. The paper does not attempt to be an exhaustive review of what is known about teaching, or of those parts of this knowledge that have been incorporated into systems, though it does provide a number of pointers to this large literature. Rather, it attempts to highlight key issues and systems, drawn largely from work at Sussex, that exemplify the complexity of this task.

The second disclaimer concerns the particular focus of the paper on key aspects of a certain kind of teaching. We have concentrated largely on systems that embody domain knowledge or skills to be learned, rather than on more open-ended systems, e.g. that facilitate dialogue between students. This appears more “teacher-centred” than “learner-centred” but is not intended to express any value judgement between these two different ways to conceptualise educational interactions.

RESTRICTED REPERTOIRE OF TEACHING ACTIONS

Versatile human teachers have an enormous repertoire of teaching actions at their disposal. These range from cases where the teacher gets the student to do almost all the work (“explain this to me...”, “solve this problem...”, “write an essay comparing...”, “choose a project to ...”) to cases where the teacher exercises more direct agency (“this is how it is done...”, “think of it this way...”, “if I were you, I would...”), via all kinds of intermediate and indirect cases, for example where the teacher organises an educational setting which facilitates several students working together in an effective way. An educational encounter, viewed as a kind of ordinary human communication, can exploit all the richness of context, modality, interaction and content of ordinary human communication as we experience it both in day to day conversations (including features such as tones of voice, irony, humour, glances, silences and so on) and in less interactive forms such as television, theatre, books, newspapers etc. Even in a distance education context, the teacher as author of the materials will be mindful of the learning situation of the students. They will probably be isolated and so materials (such as video-clips) and activities (such as self-assessment questions) will be incorporated to keep the learner motivated and self-reflective.

Ohlsson (1987) provided an excellent critique of Artificial Intelligence in Education (AIED) in terms of its historically narrow focus on modelling and diagnosis at the expense of (the harder) remedial actions and teaching. He offered an analysis of some of the many teaching operations that might be associated with teaching a procedure of some kind, see Figure 1. Note that these were specifically concerned with teaching a *procedure*, though some of these operations might also apply in teaching a *principle* or a *concept*. These operations include teaching actions associated with setting the scene as well as with indicating the nuts and bolts of the procedure. Setting the scene can involve clarification of goals as well as justification of individual steps or pointing out similarities to similar procedures already well understood by the student. It is important to reiterate that teaching can be much more than assisting the mastery of procedures, principles and concepts.

```
PRESENTING TARGET PROCEDURE

 Define terms,
   describe procedure
   prompt recall
 Demonstrate
   interactive
   prompted
   annotated
   applied
 Practice
   guided
   annotated
   corrected
   hints
   drill

PRESENTING PRECURSORS

 Priming
 Reviewing
 Marking familiar & unfamil-
iar steps

PRESENTING PURPOSES

 Giving a goal
 Criticise precursors
 Generalisation or replace-
ment of precursors

PRESENTING JUSTIFICATIONS

 Annotating
 Transparent cases
 Equivalent procedure
 Verification
   alternative
   inverse
   empirical test

DEALING WITH ERRORS

 Reveal
 Explain
 Mark

DEALING WITH SOLUTIONS

 Feedback
 Prompt self-check
 Prompt self-review
 Prompt self-annotation
```

Figure 1: Teaching a Procedure: Some Principles of Intelligent Tutoring (Adapted from Ohlsson (1987))

Perhaps the most interesting items on Ohlsson's list are those associated with reactions to the student getting his or her answer correct: actions designed to get the student to check the robustness, applicability, assumptions etc of the solution. Will it work for all cases? Supposing the initial conditions had been slightly different? Is it expressed in the most general form? Is it similar to other solutions? Are there similarities in the way the solution was constructed compared to other problem-solving episodes? Is the solution optimal? Was the problem-solving optimal? And so on. The difficulty of achieving this with a machine teacher arises not from any difficulty of posing the right kinds of question to the student, but of being able to undertake any but the most cursory analysis of the student's answer. An associated problem here is the restricted modality (typically text and diagrams on a screen) that was then available. There is something rather special about voice and gesture that can cut through complex material to reveal the essential point. The latest generation of embodied agents exploit this, as we see later.

While it may be possible to constrain the language with which the student refers to the knowledge domain itself, it is much harder to constrain the meta-language in which the student makes evaluative statements about solutions. The standard tricks at the domain level of using semantic grammars, menus or other devices to restrict input are much less easy to apply to meta-language because the domain has been broadened and the student is asked to make comparisons across the domain or indeed between domains. With notable exceptions, such as Collins and Brown (1988), the designers of systems targeted by Ohlsson were much less interest than now in the whole issue of metacognition and tools to support planning and reflection (say) were less common.

There also tended to be a basic asymmetry between student and machine teacher in that the complexity of the textual or diagrammatic output from the machine teacher was usually far in excess of the complexity of typed input from the student that could be understood by the machine teacher. Sometimes systems allowed students to type in free form text, but typically this was either ignored or only partially comprehended by the system. Of course, a student could be monitored interactively while building up a complex object (e.g. such as a Lisp program in the Lisp tutor: Anderson and Reiser, 1985) or be provided with a post hoc analysis of a complex object that he or she has built (such as a Pascal program in Proust: Johnson and Soloway, 1987), but in general machine teachers made much higher demands on the language and image understanding of their students than their students made of them. If the student thought about the domain in a different way from the machine teacher, or solved problems in an idiosyncratic manner there was usually no way that the student could tell the machine teacher that this was the case or any way for the machine teacher to be able to make an evaluative, comparative comment about the student's view or method in comparison with its own.

Although we have concentrated on Ohlsson's (1987) critique, he was not the only researcher to have called into question the teaching capability of that generation of machine teachers. In another critical analysis of Intelligent Computer Assisted Instruction (ICAI), Ridgway (1988) summed up the activities missing from machine teacher's repertoire as:

1. Encouraging explanations from pupils to each other and to the teacher.
2. Group work that facilitates metacognitive activities and is itself inherently valuable,
3. Metacognitive skills such as self-explanation and self-evaluation.

Although concerned with a wider class of systems, namely advice-giving expert systems, Carroll and McKendree (1987)[2] criticised many approaches then used for machine teachers as lacking in empirical justification or generalisability across domains. For example, with respect to Socratic Tutoring, they wrote:

[2]Many of the questions raised in this wide-ranging survey of the field are still relevant today.

> "It is likely that this is an effective style for interactive tutoring in many situations, but it is also significant that no evidence is offered to support this assumption. The possibility exists that the Socratic style is often adopted for tutorially irrelevant reasons. Giving the system control of the dialogue allows a simple question-list knowledge structure."
>
> (Carroll and McKendree, 1987, page 15)

In summary, the criticisms above can be reduced to two major issues. First, tutoring systems have focused on too narrow a range of types of educational interaction, i.e. taking rather a teacher-centred view of the enterprise and not attempting a more learner-centred facilitating role. Second, even within a teacher-centred framework, most systems have adopted rather a narrow range of teaching tactics and strategies.

So in principle how could the issues raised in these criticisms, especially the second, be dealt with?

DEVELOPMENT OF TEACHING STRATEGIES

There are three principled methodologies for developing the teaching expertise in AIED systems. First is the observation of human teachers followed by an encoding of effective examples of these teacher's expertise, typically in the form of rules. The second is based on learning theory and derives a teaching theory from that. The third is based on observations of real students or on a runnable simulation model of the student and derives a teaching theory from experiments with such students or models of students.

Derived from expert human teachers

An influential early example of the methodology of learning from expert human teachers was Socratic Tutoring (Collins et al., 1975). Socratic Tutoring provides a number of detailed teaching tactics for eliciting from and then confronting a learner with her misconceptions in some domain. A generalisation of this approach, "Inquiry Teaching" is offered by Collins and Stevens (1991). A more recent example of the general methodology is provided by Lepper et al. (1993) who analysed the methods that human teachers use to maintain students in a positive motivational state with respect to their learning.

There is an issue in using expert teachers as a source. Typically they will have been influenced to a greater or lesser extent by the theories of teaching and learning they were exposed to during training, so there is a danger that one might be simply observing these theories filtered through their application by the chosen experts. Another issue is that there are many different styles and philosophies of teaching, further fragmented by individual personality differences and domain norms. Finally what teachers say they do and what they actually do may be at odds with each other, see Bliss et al. (1996), described later.

There are many ways to examine the issue of what constitutes expert human teaching behaviour. One way is to have regard for the literature on expertise and analyse the ways that expert teachers are similar to experts in other fields. For example, Sternberg and Horvath (1995) reviewed work on the structure of expert teacher knowledge. Typically, this is found to be much more highly structured than that of novice, or even experienced teachers. They looked at the "efficiency" with which these kinds of expert solved problems in their domain. They found that, as with other experts, automaticity lends speed. They also looked at "insight" issues and concluded that expert teachers have a better insight into the deep rather than the surface structure of learning and teaching situations.

Although not concerned to assist the implementation of machine teachers, the analysis by Schoenfeld (1998) of how teachers' context-specific beliefs, goals and knowledge are activated and interact in classroom settings is couched in the language of cognitive science. The author analyses a number of mathematics lessons in enough detail to demonstrate how far we yet have to go in order to duplicate skilled human teaching and to offer hints towards a framework for such an enterprise.

Another way is to look at the fine structure of how human teachers deal with particular issues, such as maintaining students' motivation (e.g., Lepper and Chabay, 1988), offering corrective feedback (e.g., Fox, 1991), or detecting and repairing dialogue failures (Douglas, 1991). In the latter case, Douglas found in her study of teachers that "Expert and novice tutors made about the same number of [communication] failures, but the expert was markedly better at detecting and repairing them." — an observation that supports Sternberg and Horvath's (1995) analysis of teaching expertise, above.

Among the many studies cited by Sternberg and Horvath (1995) was an empirical analysis of expert and novice teachers in the area of mathematics (Leinhardt and Greeno, 1991). Echoing Ohlsson's (1987) critique, they found that

> "The expert teachers had, with the class, a large repertoire of routines, usually with several forms of each one. In some cases, we observed teachers apparently teaching new routines to their classes. The main features of these mutually known routines were that (a) they were very flexible, (b) order could be shifted and pieces taken from one segment and applied to another, (c) little or no monitoring of execution was required, and (d) little or no explanation was required for carrying them out. These routines had simple, transparent objectives: to increase the amount of time that students were directly engaged in learning or practicing mathematics and receiving feedback, to reduce the cognitive load for the teacher, and to establish a frame that permitted easy transmission of information in mutually known and recognized settings."
>
> (Leinhardt and Greeno, 1991, page 265)

More recently, in a series of studies Graesser et al. (2000) studied both expert teachers and non-expert human tutors [3]. They found that even untrained human tutors were extremely effective *and* that their methods did not seem to correspond to any of the standard methodologies such as Socratic Tutoring (Collins et al., 1975), error identification and correction (Corbett and Anderson, 1992) or sophisticated motivational techniques (Lepper et al., 1993). However they concluded that

> "Tutors clearly need to be trained how to use the sophisticated tutoring skills because they do not routinely emerge in naturalistic tutoring with untrained tutors. We believe that the most effective computer tutor will be a hybrid between naturalistic tutorial dialog and ideal pedagogical strategies"
>
> (Graesser et al., 2000, page 51)

The way that expert human teachers interweave a wide range of actions that deal with cognitive, metacognitive and affective issues is exemplified by Lajoie et al.'s (2000) observation of an expert medical instructor. They observed how the instructor assigned roles to each of the participants in the small group of first year medical students discussing patients.

> "One student was asked to present an actual patient case to the group, describing the patient's relevant medical history and current situation. A second student was

[3] The distinction between "teaching" and "tutoring" is open to debate. In referring to human teachers, we will use the term "tutor" to imply someone with less formal training in pedagogy than a "teacher". The literature of Artificial Intelligence in Education uses the terms "tutor" and "teacher" interchangeably irrespective of the expertise of the system so designated.

> then asked to summarize the same case based on the verbal account of the first student. Next, a third student was asked to produce a problem list for the patient at the blackboard with the assistance of the other students in the group. Finally, the fourth student was asked to lead the group in developing a list of differential diagnoses for the case."
>
> (Lajoie et al., 2000, page 59)

Through the organization of the student's roles and through both specific and general feedback the instructor was able to deal with a number of issues simultaneously. At the cognitive level, both the division of labour and his feedback provided scaffolding for the complex, cognitive task of arriving at a diagnosis. At the metacognitive level the interaction between students and their feedback provided models for reflective self-critical examination of how data was being used and how decisions were being arrived at. At the affective level, encouragement was being provided when needed. And finally, at the "community of practice" level, students were being apprenticed into medically accepted ways of behaving.

It is unrealistic to expect that an all-embracing, prescriptive theory of teaching will easily emerge given the complex, social nature of the enterprise. It would be like expecting a prescriptive theory of "being a politician" or "being an actor". Of course in each of these activities there are guidelines which the novice teacher (or politician or actor) can make use of and some theories and practical tips on, say, how to be convincing, how to explain effectively or how to reflect on performance. For example, in the latter case, it can be very useful for a teacher to see a videotape of a lesson he or she has taught and to discuss the performance with a more experienced person.

But these theories can never be entirely prescriptive in that the activities do not occur in a vacuum but often depend for their effectiveness on the personalities of the participants. An authoritarian, disciplined teacher may be just as effective (along certain dimensions) as a more easy-going person with a laissez-faire approach. It depends on how well the individual teacher can exploit his or her own personality traits for the job in hand (Rutter et al., 1979). The above should not be taken to mean that the theoretical study of teaching interactions is misguided, only that social phenomena are enormously subtle.

Within the overall space of possible teaching situations, let us examine two in particular: dealing with motivational issues and dealing with errors. These two have been chosen because they interact with each other and also because they nicely illustrate a polarisation of emphasis between systems that Ohlsson (1987) criticised and what has been achieved since then. Of course, the very notion of "dealing with errors" betrays an expectation of correct and incorrect answers, and this will apply only in a limited kind of learning situation.

In order to maintain the sense of historic continuity, the issue of dealing with errors is taken first. This issue has been central in the development of machine teachers.

Dealing with errors

Many machine teachers have addressed the issue of dealing with student errors, so it is natural that designers of such systems have looked to human expert teachers for insight. An important teaching controversy around such systems is the issue of the tactical *vs.* the strategic value of providing immediate help when errors are detected. Tactically the student is helped at the point where they have the best chance of understanding and exploiting that help. Strategically, however, they may be being denied opportunities to figure out for themselves what went wrong and what to do next.

A review of the literature on the comparison between expert human teachers and these kinds of intelligent tutoring systems is provided by Merrill et al. (1992b). They noted that different experts adopted different styles of tutorial feedback:

> "Fox (1991) and Lepper et al. (1991) argued that tutors use very subtle feedback upon errors or obstacles to maximise students' problem solving success. The McArthur et al. (1990) results also suggest that tutors follow students' solutions very carefully, but indicate that this feedback can be very directive. McArthur et al. argued that tutors give explicit feedback, sometimes even telling students how to solve a problem, and carefully structure students' tasks by reminding them of the problem goals. Littman et al. (1990) and Merrill et al. (1992a) argued that the context of the error is critical in determining feedback"
>
> (Merrill et al., 1992b, page 283)

There are clearly differing styles of teaching that position themselves at different points in the trade-off between providing too much help (thus potentially inhibiting the development of problem-solving strategies), and providing too little help (thus risking the possibility that students become lost and demotivated). For example, in the context of industrial training, the Recovery Boiler Tutor offered the trainee a number of "precautionary messages . . . when a full-scale disaster is imminent" rather than specifically negative feedback statements. These might redirect the student: *Have you considered* . . . ; or, draw their attention to an unobserved relationship: *Did you notice the relationship between* . . . ; or, confirm those actions which are helpful (Woolf, 1988).

An important facet of the issue of how much specific help to offer is the degree to which students are in fact sufficiently self-aware to know when they do in fact need help (see e.g., Aleven and Koedinger, 2000).

Merrill et al. (1992b) provide a critical comparison of a range of model-based tutoring systems in relation to the above findings from human tutors. An interesting issue that is explored as part of this discussion is the way that the interface to the system can be a crucial element in helping the student understand the nature of the problem-domain, over and beyond any feedback that a tutor might additionally provide (Reiser et al., 1992).

In comparing human and computer-based tutors' ways of dealing with errors, Merrill et al. (1992b) note some similarities, e.g. both help "students detect and repair errors and overcome impasses". However, they also note a number of differences. Human tutors offer less "explicit verbalizations of the student's misconception" than machine teachers; they are more flexible in the timing and the nature of their feedback; and they are more subtle in indicating to the student that an error has occurred, e.g. via slight pauses or intakes of breath.

Motivating students

An important aspect of human teaching expertise that figures only weakly in Ohlsson's (1987) criticism centres around skills employed by teachers to build and maintain students' engagement with the task and their motivation to learn.

According to Lepper et al. (1993) the focus for many expert teachers is not just on cognitive issues, such as what task to teach next or what error the student has committed. They also focus strongly on affective issues. How can the student be stimulated and challenged? How can the student's confidence and sense of control over the learning situation be maintained or improved. In pursuit of these affective goals, human teachers are often indirect in their way of helping students to set goals or in reacting to their errors: "Now tell me *how* you got that 6?".

In the tricky issue of detecting the student's current motivational state Lepper et al. (1993) suggest that experienced teachers make use of "the student's facial expressions, body language, intonation, and other paralinguistic cues".

There is another way of gathering evidence about the students motivational state. Student effort, rather than performance, is also a reasonably reliable indication of intrinsic motivation (Keller, 1983). Learners who display a high level of effort (detected through their activities, suggestions, responses) may deserve praise even when their performance is non-optimal. There is a wide literature on the relation between extrinsic rewards (such as praise) and intrinsic motivation. Eisenberger's (1992) work suggests rewarding effort is effective over the long-term and an extensive meta-analysis of the literature by Cameron and Pierce (1994) suggests that (contrary to received wisdom) extrinsic rewards such as praise do not decrease intrinsic motivation.

Observation of human teachers thus lends force to Ohlsson's (1987) criticisms. Indeed, they go beyond it by pointing out the importance of affective issues in determining how expert teachers behave. Before turning to examine systems that post-date Ohlsson's criticisms, we briefly examine the two other methodologies for deriving the behaviour of a machine teacher.

Derived from learning theory

We have already noted that there are three methodologies for deriving a teaching theory for a machine teacher. The second of the three methodologies starts from a learning theory and derives appropriate teaching tactics and strategies from that theory.

Before continuing with our discussions about how learning theories have informed the design of computer systems that teach we need to say a few words about the nature of knowledge itself. The nature of knowledge has long been an area of active research and discussion amongst philosophers, psychologists and educationalists as well as those involved in the development of artificial intelligence systems. Is knowledge absolute or relative? Does it exist as an external object that can be known or is it bound up with each individual's environment and experience? What are the implications of what we belief about the nature of knowledge for the way we perceive our own knowledge and the processes by which we acquire that knowledge? In other words how do our beliefs about knowledge interact with the way we understand the process of learning? There is no room for us to do justice to a discussion of the nature of epistemological beliefs and their role in learning here. We therefore raise the question in the reader's mind as one that they would need to consider for a fuller exploration of the nature of learning theories. However, in this current paper we now place ourselves one step removed from this discussion and concentrate on the nature of the learning theories that have influenced system design rather than the epistemology upon which they are founded. For examples of work in this area the reader is referred to von Glasersfeld (1984, 1987); Wilson and Cole (1991); Wilson (1997).

Let us now look at some specific examples of learning theories and their interactions with teaching theories. Conversation Theory (Pask and Scott, 1975) and its reification in various teaching systems is an example of this approach. As with Socratic Tutoring, Conversation Theory is concerned essentially with epistemology rather than with affective aspects of teaching and learning. It is based on a view of learning consisting of two interacting processes. One operates at the domain level, for instance adding links, facts, rules and principles. The other works at the meta-level, noting gaps and inconsistencies in what is known at the domain level. These two processes can operate inside an individual learner or they can be distributed as roles between more than one participant. Overall the theory sets conditions to ensure that the learner constructs a multifaceted understanding of a domain that allows her to describe (to herself or to others) the inter-relationships between concepts. In some ways this is echoed by the "self-explanation" view of effective learning (Chi et al., 1989). A further example of the second methodology, which also partially addresses some of the affective issues, is Contingent Teaching (Wood and Middleton, 1975). Here the idea is to maintain the learner's agency in a learning interaction by providing only sufficient assistance at any point to enable her to make progress on the task. The evaluation of this strategy in the hands of non-teachers who had been deliberately taught it shows that it is effective. However it was found sometimes to go against the grain for experienced teachers who often wish to provide more help at various points than the theory permits (Wood et al., 1978). Successfully helping teachers to apply a teaching strategy based on scaffolding (Wood et al., 1976) can be difficult, as the study by Bliss et al. (1996) shows. They found that, even after teachers' reflective observation of their own and their colleagues lessons, focusing on opportunities and methods of scaffolding pupils' learning:

> "...they professed improved practice and demonstrated greater confidence in discussing scaffolding, but there was no significant increase in the number of instances that could be described as scaffolding. When scaffolds were used these were usually on a one-to-one basis. It was during this phase that we realised that our teachers could 'talk scaffolding' but appeared to implement it only marginally. Their focus was on teaching rather than on pupils' learning."
>
> (Bliss et al., 1996, pages 44-45)

The range of tutors developed by Anderson and his colleagues has provided an influential, and controversial, model of teaching in the area of dealing with errors, see for example, (Anderson and Reiser, 1985; Anderson et al., 1985; Corbett and Anderson, 1992; Anderson et al., 1995), and more recently (Koedinger et al., 1997, 2000). The form of teaching has been characterised as one-to-one with fine-grained diagnosis and remediation for multi-step problem-solving in a variety of formal domains such as geometry, programming and algebra. It applies only to problems with tightly constrained solution methods and this is clearly limiting in terms of the kinds of educational interaction between machine teacher and student that can be supported. However, recent work by Koedinger et al. (1997) using the PAT system to teach algebra has shown how some of these limitations can be circumvented by paying special attention to contextual factors (e.g. as emphasised by Schoenfeld, 1998). In particular, great care was taken to involve the schools and the teachers who would be using the system and careful thought was given to the use of the Tutor within the classroom. The system was used not on a one-to-one basis but by teams of students who were also expected to carry out activities related to the use of the tutor, but not involving the tutor, such as making presentations to their peers. This use of explanation between learners helps to counter one of Ridgway's (1988) criticisms, mentioned earlier.

The design of these systems is derived from a theory of learning that, at base, provides an account of the development of expertise which explains how declarative knowledge is transformed into procedural knowledge and how this transformation can be supported by learning environments (Anderson, 1990). A consequence of the theory is that attention is paid to ensuring that learners are kept aware of the goal and sub-goal hierarchy of the problem solving they have embarked on.

The intelligence of these systems is deployed in several ways. Model Tracing, based on representing knowledge of how to do the task in terms of production-rules, is used to keep close track of all the student's actions as the problem is solved and flag errors as they occur, such as misplotting a point or entering a value in an incorrect cell in the spreadsheet. It also adjusts the help feedback according to the specific problem-solving context in which it is requested. Knowledge Tracing is used to choose the next appropriate problem so as to move the students in a timely but effective manner through the curriculum.

Learning theories are still being used to inform system design: for example, Constructivism (Akhras and Self, 2000) and Reciprocal Teaching (Chan and Chou, 1997). In addition, Grandbastien (1999) stresses the need for effective methods to access, organise and use the expertise of the teacher or trainer. Starting from a model of learning, Winne (1997) suggests how students might be helped to develop better "self-regulated" learning capability (i.e. improve their metacognitive skills).

Derived from studies of students

The third methodology for deriving a teaching theory is based on observation of students. One methodology observes how students of different types respond to a particular teaching method, for example assessing how students of differing ability fare with a particular machine teacher.

Another methodology compares differing methods across students, and a third methodology observes interaction effects between student characteristics and teaching methods.

These methodologies come in two forms. There can be empirical observations of *real* students or there can be analyses of the reactions of *simulated* students.

Studies of Real Students

Within the educational literature as a whole there is a huge literature on how students of differing characteristics respond to differing teaching methods. The range of characteristics include gender, ability, learning style, background knowledge, age and so on. For a review of this huge area the reader is referred to Cronbach and Snow (1977). Within the artificial intelligence in education community, many studies have looked at how students of differing ability and background respond to particular systems. As a single recent example of this style of work we select Arroyo et al. (2000). They categorised a cohort of students by gender and by level of cognitive development. They wanted to establish how variations in the style of hints in the context of an arithmetic program interacted with gender and with cognitive development. Hints varied on two dimensions: degree of interactivity and the nature of the symbolism used. They looked at the reduction in the number of mistakes on a problem following a hint as one of the dependent variables. They found a number of interaction effects (e.g. that "high cognitive ability students do better with highly symbolic hints while low cognitive ability students do worse with highly symbolic hints"). These and related results should enable the program to make "macroadaptive" (Shute, 1995) changes to its teaching strategy to suit particular sub-groups of students.

Despite the huge wealth of work, it is difficult to derive general guidelines about the differential effect of students' characteristics of sufficient precision and reliably to support the design of machine teachers.

In terms of looking at the effects of variations of teaching method, an important indirect influence on progress has been the work of Bloom (1984) and his colleagues. They investigated how various factors, such as cues and explanations, reinforcement and feedback, affect student learning, taking conventional classroom teaching as the baseline. They found that highly individualised expert teaching, shifts the distribution of achievement scores of students by about two standard deviations compared to the more usual situation where one teacher deals with a classroom of students. They also found that the range of individual differences reduced.

This two standard deviation improvement, or Two Sigma shift, has become a goal at which designers of machine teachers aim. A standard method of evaluation of such a system is to compare it with conventional non-computer-based interaction teaching the same topic, though there have been some comparisons of "smart" and "dumb" versions of the same software. For more on evaluation of AIED systems, see du Boulay (2000); Self (1993).

Studies of Simulated Students

The second form of the methodology that is based on observation of students uses simulated rather than real students. The examples we know of this methodology typically compare different teaching methods across identical simulated students rather than modelling students of differing characteristics and observing how a particular teaching method affects them differentially. This methodology builds a computational model of the learner or of the learning process and derives a teaching strategy or constraints on teaching behaviour by observing the model's response to different teaching actions. For example, VanLehn et al. (1994) compared two strategies for teaching subtraction to a production rule model of a subtraction learner and concluded, on the basis of the amount of processing engaged in by the model, that the "equal additions" strategy was more effective than the more widely taught "decomposition" strategy. With a similar general methodology VanLehn (1987) derived "felicity conditions" for the structure of tutorial examples, for instance that they should only contain one new subprocedure.

This methodology offers an interesting avenue for research but in terms of making predictions about how real students will react, it depends crucially on the fidelity of the underlying simulations.

EXAMPLES OF RECENT PROGRESS

This section does not aim to be a comprehensive review of the current state of the art in modelling teaching for tutoring systems. It offers a number of examples of recent and relatively recent systems that attempt to go beyond the restrictions outlined by Ohlsson (1987) and Ridgway (1988). Two of the examples derive directly from the analysis in the previous section. Thus the first subsection exploits the literature on expert teachers to tackle the central issue of the affective dimension in teaching and describes a system concerned with modelling teachers' motivational expertise. The second subsection shows how Contingent Teaching and a Vygotskian learning theory can be exploited and deals with an aspect that has been of central interest from the start of AIED, namely adjusting the kind of activities and the help provided to students to succeed on those activities. The system can dynamically adjust the terminology it uses to describe its domain to the student as well as make adjustments to both the complexity of the tasks it sets as well as the help it provides.

Because of their increasing visibility, the third subsection looks at several examples of pedagogical agents and examines how they are affecting the debate about modelling teachers' behaviour, including the issue of the perceived plausibility of such systems.

Note that one way to track the changes in the teaching ability of modern systems is to examine tools for building machine teachers, namely authoring systems. This is not the place to survey such systems and an excellent recent survey is provided by Murray (1999). He is upbeat about their capability to represent tutorial strategies and tactics beyond simple issues of curriculum sequencing and planning:

> "Instructional decisions at the micro level include when and how to give explanations, summaries, examples, and analogies; what type of hinting and feedback to give; and what type of questions and exercises to offer the student. ... Also characteristic to systems in this category is the ability to represent multiple tutoring strategies and "meta-strategies" that select the appropriate tutoring strategy for a given situation."
>
> (Murray, 1999, page 102)

Among the many systems surveyed by Murray (1999), REDEEM stands as a good example of an authoring system with the capability of specifying a wide variety of teaching strategies (Major et al., 1997). This system provides tools for authors (usually teachers) to reuse and reorganize existing non-adaptive pages of tutorial material into a responsive and adaptive system.

Adding motivational competence

As we have seen, theories of instructional motivation elaborate the influence of issues like confidence, challenge, control and curiosity in the learning process (Keller, 1983; Malone and Lepper, 1987) and suggest instructional tactics to keep the student in an optimal learning state and provide appealing and effective interactions. Of course, it is an open question as to how far in practice one needs a separate theory of motivation. It could be argued that if one gets the cognitive, metacognitive and contextual issues right, then all will be well. Each of these is itself a complex issue, and so for the purposes of progress it seems sensible to clarify the means by which students can be motivated.

Work on motivational issues is proceeding along two fronts. The first reasons about the affective state of the *teacher*. With the rise in interest in creating pedagogical agents, increasing attention is being paid to equipping them with affective competence. For example, Lester et al. (1999) describe the techniques underpinning COSMO, a pedagogical agent that can adapt both its facial expression, its tone of voice, its gestures and the structure of its utterances to indicate its own affective state and to add affective force and focus to its interactions with learners. By presenting its reactions to the student in a more varied, human-like way the hope is that students will be better motivated to learn and better able to judge what is important in what is being advised. These issues are elaborated in a later section of the paper.

On another front work is progressing on a system, MORE, to reason about the affective state of the *student* (del Soldato, 1994; del Soldato and du Boulay, 1995). To this end implementing motivational techniques demands shaping the system, including domain representation and student model, in several aspects. In what follows the assumption is that the student is working on topics assigned as part of a curriculum rather than working on topics purely of their own choosing, in which case the motivation issues are likely to be very different. In particular, the system must:

1. Detect the student's motivational state;
2. React with the purpose of motivating distracted, less confident or discontented students, or sustaining the disposition of already motivated students.

The notion of a system's reaction — triggering particular motivational tactics — suggests that a comprehensive instructional plan should consist of a "traditional" instructional plan combined with a motivational plan. Wasson (1990) proposed the division of instructional planning into two streams: content planning ("which topic to teach next"), followed by delivery planning ("how to teach the next topic"). At first sight the motivational plan seems to be completely embedded in the delivery plan. However, motivational tactics do not always simply complete the traditional content planning: sometimes they compete with it as well. A typical example of such a conflict is the necessity for less confident students to build their confidence by accumulating experience of success, in which case the system could provide problems likely to be positively answered — based on topics that the student *already* knows. This is related to the need for practice in learning, an issue not well explored in AIED systems. Furthermore, while the detection of a learner's motivational profile shapes the student model, the system's reaction (e.g. suggesting an easier problem, asking a puzzling question) depends on the resources available in the domain representation.

Typical domain-based planners select actions according to whether the learner knows a topic or has mastered a skill. The methodology is twofold: detecting the current state of the learner's knowledge and skill (student modelling) and reacting appropriately in order to increase this knowledge and skill (teaching expertise). To take account of motivational factors, we have extended the twin activities of "detecting the state" and "reacting appropriately" by adding a *motivational state* and *motivational planning* to the traditional ITS architecture. Sometimes the advice offered by the motivational planner disagrees with the domain-based plan, while in other cases both plans complement each other (del Soldato and du Boulay, 1995).

Let us consider, as an example, the situation in which the student succeeds in solving a problem, in this case finding the bug in a program. A typical domain-based planner would acknowledge the right answer and suggest (or directly provide) a harder problem, thus making sure the student is traversing the domain in a progressive manner. Such behaviour is well exemplified by Peachey and McCalla's (1986) instructional planner: when the learner masters an instructional goal, the planner focuses next on goals that require the topic just mastered as pre-requisites, traversing the domain in the direction of a specific ultimate goal. In this case, *knowing* or *not knowing* the topic, or exhibiting or not exhibiting the relevant skill, is the only issue in the student model that drives the selection of suitable actions.

Motivational planning takes into account other variables in the student model and widens the tutor's space of possible reactions. Just by considering binary states of *effort* (little/large) and *confidence* (low/ok) results in four different situations, each one requiring a suitable set of actions from the tutor, the binary states of *effort* and *confidence.*

When the student's confidence is diagnosed as being low, the major goal for the planner is to help the learner regain a reasonable level of confidence, and one of the tactics for improving confidence is to increase the student's *experience of success.* The tutor should then select a task likely to be solved successfully again (e.g. a similar task to the one the student has just accomplished). This is a clear example of disagreement between the domain-based and the motivational planner, since simply traversing the domain to the next harder topic has been deliberately avoided.

On the other hand, if providing a right answer requires little effort from the student (even an insecure one) the tutor should move to harder tasks. The tutor should make the *difficulty-level promotion* very clear, both by praising the successes obtained so far and warning about the new difficulties which are likely to be encountered at the harder level. The student then is prepared to cope with new failures without feeling too de-motivated.

Let us now consider the case of a task that does not require very much effort from a normally confident learner. For a typical domain-based planner such a situation would be ideal, whereas from a motivational perspective the task could be perceived as being irrelevant or "boring", or in other words, de-motivating. The tutor should then increase the degree of challenge provided by the interaction, by adjusting the difficulty level to a harder one where the student would not always (easily) perform the task, and some effort would have to be spent to achieve success.

Similar analyses have been made for cases where the student fails at a task or gives up on a task and these tactics were implemented as production rules in the motivational planner. The issue here is not so much whether these particular tactics are correct, but the fact that tactics such as this can be modelled explicitly within the system. If these are not the best rules, then others can replace them without having to redesign the system from scratch.

The need for flexibility in the way that motivational tactics are implement is underlined by the evaluation of MORE. One of the issues that emerged was the reluctance of some students to accept certain teaching tactics from *a machine* as opposed to a human — the refusal to provide help when asked or the refusal to allow the student to give up on an unsolved problem. This is an example of the "plausibility problem", which we discuss later.

Judging task difficulty and degree of assistance

Assuming that a learner is in a reasonable state of motivation, the teacher can then focus on what the learner should do and how they should be helped.

We now discuss the issue of adjusting the complexity of the learning environment, the content and difficulty of the activities, the language in which they are expressed, and the quality of hints and suggestions in interactive learning environments (ILEs). In particular, we describe the educational philosophy underpinning the Ecolab software.

The Ecolab is an Interactive Learning Environment (ILE) which aims to help children aged 10–11 years learn about food chains and webs. The Ecolab provides a flexible environment which can be viewed from different perspectives and run in different modes and in increasingly complex phases. In addition to providing the child with the facilities to build, activate and observe a simulated ecological community, the Ecolab also provides the child with small activities of different types, such as finding out what animals eat, which are predators and which are prey, establishing the energy changes associated with feeding, and setting up a self-sustaining small ecosystem. The activities are designed to structure the child's interactions with the system. They provide a goal towards which the child's actions can be directed and vary in the complexity of the relationships which the child is required to investigate.

This system explores the way that Vygotsky's Zone of Proximal Development can be used in the design of learner models (Luckin, 1998; Luckin and du Boulay, 1999). The theoretical foundation requires the system to adopt the role of a more able assistant for a learner. It must provide appropriately challenging activities and the right quantity and quality of assistance. The learner model must track both the learner's capability and her potential in order to maintain the appropriate degree of collaborative assistance.

One of the links between the work on the ZPD and the work on motivation is the notion of *effort*. For the motivational planner the amount of effort expended by the student is a measure of her motivational state. For Vygotsky, an appropriate degree of mental effort is a pre-requisite for learning.

The Zone of Proximal Development (ZPD) (Vygotsky, 1978) is created when two or more people form a collaborative learning partnership in which the more able members enable the less able members to achieve their goal. In order for a collaborator to be successful in the role of a more able learning partner she must construct a shared situation definition (Wertsch, 1984) where all members have some common knowledge about the current problem. This intersubjectivity can only be achieved if the teacher/collaborator has a dynamic representation of the learner's current knowledge and understanding. The ZPD also has a spatial analogy which quantifies a learner's potential (Vygotsky, 1986). It is the fertile area between what she can achieve independently and what she can achieve with assistance from another. In essence the ZPD requires collaboration or assistance for a learner from another more able partner. The activities which form a part of the child's effective education must be (just) beyond the range of her independent ability. The learning partner must provide appropriately challenging activities and the right quantity and quality of assistance. In the Ecolab the learning partner role is adopted by the system, and so the learner model must track both the learner's capability and her potential in order to maintain the appropriate degree of collaborative assistance.

The strong focus on adapting to the user by adjusting the amount of help that is initially offered is similar to the adaptive mechanisms in the SHERLOCK tutors (see e.g., Katz et al., 1992; Lesgold et al., 1992). A difference from SHERLOCK is that there is also adjustment both to the nature of the activities undertaken by users and to the *language* in which these activities are expressed. The working assumption is that more abstract language is harder and learners move from the concrete toward the abstract. An alternative view might offer the abstract terminology earlier as an aid to generalisation. The emphasis which the Ecolab places upon extending the learner beyond what she can achieve alone and then providing sufficient assistance to ensure that she does not fail also sets it apart from other system's such as that of Beck et al. (1997), which generate problems of controlled difficulty and aim to tailor the hints and help the system offers to the individual's particular needs. The Ecolab extends the work done with other systems which have used the ZPD concept in relation to the learner modelling task such as the system of Gegg-Harrison (1992) which offers the learner guided problem-solving sessions in which they are given assistance in solving difficult Prolog problems.

The Ecolab can assist the child in several ways. First, it can offer 5 levels of graded help specific to the particular situation; second, the difficulty level of the activity itself can also be adjusted (activity differentiation). Finally, the definition of the domain itself allows topics to be addressed by the learner at varying levels of complexity and (independently) using terminology of varying levels of abstractness. So, for example, activities can involve simple bilateral relationships between say "rabbits" and "grass", or the same simple relationship described in the more abstract terms "herbivore" and "primary consumer". In addition, more complex relationships (such as between distant members of the same food web) can also be described either in simple or more abstract language.

The Ecolab uses a bayesian belief network model of the difficulty of transitions between nodes and the history of success and help required at previous nodes to decide:

- Which node in the curriculum will be tackled next — which level of complexity and which level of terminology abstraction.
- What level of help will be offered.
- How much activity differentiation will be offered to the child.

The system maintains a model of the ecology curriculum based on a Bayesian Belief Network. Each node in the curriculum represents a rule to be learned. The rules are linked via pre-requisites that impose a partial order on the rules in the curriculum. There is a starting node and also a notional finishing node, namely the most complex rule explored via the most abstract terminology.

Each node is associated with a probability value that indicates the likelihood that the learner can complete, unassisted, activities associated with that node. The system uses these values to distinguish nodes that are either too easy (outside the ZPD), just too hard (within the ZPD) or much too hard (outside the ZPD) for that learner to complete unassisted.

The next node for the learner is chosen from those that are just too hard. The system uses data about the learner's progress with previous nodes to set both the degree of difficulty of the activity chosen as well as the quality of the initial help which is offered if needed.

Once an activity is completed the actual amount of help that the learner used is noted. This may be more than was expected, if the activity turned out to be harder, or can be less than expected if in fact the learner did not need any help. The amount of help actually provided is used by the system to update the probability value of mastery at that node. This value is then propagated through the network to update the probability values at all other nodes linked to it via pre-requisites (Luckin, 1998). A node is again chosen that is just too hard. This may involve either a progression through the curriculum (i.e. to a more complex rule or towards more abstract terminology) or staying at the current node and tackling a different activity.

Student's input to the Ecolab was largely unambiguous, e.g. button presses to choose different aspects of the interface or to choose animals and their actions to assemble into small programs. Where inappropriate actions were chosen the system generated help messages at the appropriate level of specificity, depending on its view of the learner's degree of mastery of the topic. When the learner made such a mistake the Ecolab did not try to reason about what the learner might have had in mind. Its model of the learner was an overlay of standard ecology knowledge expressed as a set of probabilities that the learner had mastered each of the topics.

An issue that emerged from the work with Ecolab (which is the subject of current research) concerns the pupils' degree of insight into their own state of learning and possible need for help, see also (Aleven and Koedinger, 2000) mentioned earlier.

Making the teacher manifest and believable

The generation of systems critiqued by Ohlsson (1987) realised their teaching expertise through their textual interactions with students or through changes to the interfaces onto the domains being studied. With the rapid improvement in graphical and audio technology many new possibilities for animated pedagogical agents present themselves. Such systems still have to address the same range of teaching problems as before, but they can now bring a wider range of tactics to bear (e.g. a change of facial expression, or a change of verbal emphasis).

Johnson et al. (2000) describe a number of pedagogical agents for different domains, including Steve (for teaching about operating machinery), Adele (for teaching medicine), Herman the Bug (for biology) and COSMO (for advice about internet protocols). They argue that such systems bring extra possibilities in the following areas:

- Interactive demonstrations: where the agent can do the task, point to items in a simulated environment as well as hand over the task to the learners and comment on their performance.

- Navigational guidance: where the agent can assist the learners to find their way round and establish their bearings within a complex VR world.

- Use of gaze and gesture: where the agent can exploit its own gaze and gesture to indicate its current focus of attention in a non-verbal manner.

- Use of non-verbal feedback and conversational signals: where the agent can indicate that the student has completed a task correctly or incorrectly by different kinds of nod of the head or by changing its facial expression, or exploit eye contact, or adjust tone of voice and emphasis. These kinds of capability go some way towards providing both the subtle and non-intrusive feedback employed by expert human teachers, described earlier.

- Conveying and eliciting emotion: where the agent can indicate surprise, pleasure, displeasure, puzzlement and other emotions appropriate to the current state of the learning interaction.

- Virtual teammates: where the agent can play the role of one or more teammates in tasks where the learner needs to learn how to coordinate actions for a role with the actions taken by other role players.

Embodied pedagogical agents offer the possibility that some of the subtle techniques employed by expert teachers (Merrill et al., 1992b) can now be applied by machine teachers. Of course, these extra possibilities bring extra complexity: for example, not just a matter of deciding what to say and when to say it, but also a matter of exactly how to say it. So one of the central and long-standing problems of the field has re-emerged with new force.

In addition to any problems of educational effectiveness in practice, machine teachers are vulnerable to what Lepper et al. (1993) call the "Plausibility Problem":

> "Even if the computer could accurately diagnose the student's affective state and even if the computer could respond to that state (in combination with its diagnosis of the learner's cognitive state) exactly as a human tutor would, there remains one final potential difficulty: the plausibility, or perhaps the acceptability, problem. The issue here is whether the same actions and the same statements that human tutors use will have the same effect if delivered instead by a computer, even a computer with a virtually human voice." (Lepper et al., 1993, page 102)

In other words, will human teaching tactics and strategies, or tactics derived from learning theories or learning systems work effectively for a machine teacher? We already noted how students found certain actions of the MORE machine teacher unacceptable. For a more extended discussion of this issue see Lepper et al. (1993); du Boulay et al. (1999).

CONCLUSIONS

How far has the situation improved from that described by Ohlsson? Since the mid-eighties there have been two very useful developments. Partly as a result of the desire to improve the capabilities of such systems, there has been an increasing amount of research into human expert teaching practice. Of course, teaching has been studied for millenia, but the more recent work has studied it at a level of granularity and with the possibility that the tactics and strategies observed might be implementable. This has lead to a gradual filling in of the jigsaw of capabilities, taking a wider range of issues into account such as motivation and individual differences.

Second, the advent of pedagogical agents has again thrown the spotlight back onto the whole issue of teaching expertise and the subtlety of learner-teacher interactions. There are some encouraging evaluations of such systems (see e.g., Lester et al., 1997) but they also raise many interesting and, as yet, unresolved issues such as their plausibility and their acceptability across a range of educational contexts.

The Internet has been a huge influence on the development of systems for education. Overall this has favoured more learner-centred approaches than those in which teaching tactics and strategies are to the fore, see Collins et al. (2000); McCalla (2000) for analyses of these trends. Even within a web-based learned-centred paradigm, the system can make various automatic adjustments, e.g. to which pages are made accessible to a particular learner, see Brusilovsky et al. (1998) for an example. The introduction of networked technologies which allow learners to interact across widely distributed geographical locations enables interactions between human learners and teachers which were previously unavailable. Are the issues which were pertinent to traditional face-to-face human teaching and learning still pertinent or should we be exploring the changes in human teaching within this paradigm in order to inform our designs for intelligent systems to support this learning?

We should not lose sight of the strengths of machine teachers, despite their failings. In addition to being able to reify the learning domain and the learning or problem-solving process in ways not easily open to human teachers, machine teachers have the ability to act in a patient and *consistent* manner. This consistency can be both in terms of their knowledge and strategy as well as in terms of their emotional reactions. As human teachers we are well aware of the occasional emotional intensity of certain educational interactions and we have already cited the study by Bliss et al. (1996) that observed a disparity between what teachers said they were doing and what they actually did in the classroom. Machine teachers do not need to be prey to these problems — unless, of course, our theory of education suggests such intensity or unpredictability needs to play a role!

In future, as machine teachers evolve, no doubt we will see the emergence of personality types amongst them, with some being jokey, alert and quick-fire while others are more well-mannered and pedestrian. Each kind may suit some types of student on some occasions. As systems become more versatile we may see the emergence of the possibility of some negotiation over what is to be learned: this would be likely to help the motivation issues mentioned earlier.

We are also starting to see the emergence of systems that monitor the interactions amongst students while they learn in order to ensure that all parties play an effective role. A certain amount can be achieved here without the need for complex natural language processing techniques (see e.g., Soller, 2001).

Acknowledgements

We thank the reviewers of the first draft of this paper for their many excellent and helpful comments.

References

Akhras, F. N. and Self, J. A. (2000). System intelligence in constructivist learning. *International Journal of Artificial Intelligence in Education*, 11. to appear.

Aleven, V. and Koedinger, K. R. (2000). Limitations of student control: Do students know when they need help? In *Intelligent Tutoring Systems: 5th International Conference, ITS2000, Montreal*, number 1839 in Lecture Notes in Computer Science, pages 292–303. Springer, Berlin.

Anderson, J. R. (1990). *The Adaptive Character of Thought*. Lawrence Erlbaum Associates, Hillsdale, New Jersey.

Anderson, J. R., Boyle, C. F., and Yost, G. (1985). The geometry tutor. In *Proceedings of IJCAI'85, Los Angeles, California*, pages 1–7.

Anderson, J. R., Corbett, A. T., Koedinger, K. R., and Pelletier, R. (1995). Cognitive tutors: Lessons learned. *The Journal of the Learning Sciences*, 4(2):167–207.

Anderson, J. R. and Reiser, B. J. (1985). The LISP tutor. *BYTE*, 10(4):159–175.

Arroyo, I., Beck, J. E., Woolf, B. P., Beal, C. R., and Schultz, K. (2000). Macroadapting animalwatch to gender and cognitive differences with respect to hint interactivity and symbolism. In *Intelligent Tutoring Systems: 5th International Conference, ITS2000, Montreal*, number 1839 in Lecture Notes in Computer Science, pages 574–583. Springer, Berlin.

Beck, J., Stern, M., and Woolf, B. P. (1997). Using the student model to control problem difficulty. In Jameson, A., Paris, C., and Tasso, C., editors, *User Modeling: Proceedings of Sixth International Conference, UM97*, pages 278–288, New York. Springer Wien.

Bliss, J., Askew, M., and Macrae, S. (1996). Effective teaching and learning: Scaffolding revisited. *Oxford Review of Education*, 22(1):37–61.

Bloom, B. S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, 13(6):4–16.

Brown, J. S. and Burton, R. R. (1975). Multiple representations of knowledge for tutorial reasoning. In Bobrow, D. G. and Collins, A., editors, *Representation and Understanding*, pages 311–349. Academic Press, New York.

Brusilovsky, P., Kobsa, A., and Vassileva, J., editors (1998). *Adaptive Hypertext and Hypermedia*. Kluwer Academic Publishers, Dordrecht.

Burton, R. R. (1982). Diagnosing bugs in a simple procedural skill. In Sleeman, D. and Brown, J. S., editors, *Intelligent Tutoring Systems*, pages 157–183. Academic Press.

Burton, R. R. and Brown, J. S. (1977). Semantic grammar: A technique for constructing natural language interfaces to instructional systems. Technical Report 3587, Bolt Beranek and Newman Inc.

Cameron, J. and Pierce, W. D. (1994). Reinforcement, reward, and intrinsic motivation: A meta-analysis. *Review of Educational Research*, 64(3):363–423.

Carroll, J. and McKendree, J. (1987). Interface design issues for advice-giving expert systems. *Communications of the ACM*, 30(1):14–31.

Chan, T.-W. and Chou, C.-Y. (1997). Exploring the design of computer supports for reciprocal tutoring. *International Journal of Artificial Intelligence in Education*, 8(1):1–29.

Chi, M., Bassok, M., Lewis, M., Reimann, P., and Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, 13:145–182.

Clancey, W. J. (1982). Tutoring rules for guiding a case method dialogue. In Sleeman, D. and Brown, J. S., editors, *Intelligent Tutoring Systems*, pages 201–225. Academic Press.

Collins, A. and Brown, J. S. (1988). The computer as a tool for learning through reflection. In Mandl, H. and Lesgold, A., editors, *Learning Issues for Intelligent Tutoring Systems*, pages 1–18. Springer-Verlag, New York.

Collins, A., Neville, P., and Bielaczyc, K. (2000). The role of different media in designing environments. *International Journal of Artificial Intelligence in Education*, 11(2):144–162.

Collins, A. and Stevens, A. L. (1991). A cognitive theory of inquiry teaching. In Goodyear, P., editor, *Teaching Knowledge and Intelligent Tutoring*, pages 203–230. Ablex Publishing Corporation, Norwood, New Jersey.

Collins, A., Warnock, E. H., Aiello, N., and Miller, M. L. (1975). Reasoning from incomplete knowledge. In Bobrow, D. G. and Collins, A., editors, *Representation and Understanding*, pages 383–415. Academic Press, New York.

Corbett, A. T. and Anderson, J. R. (1992). LISP intelligent tutoring system: Research in skill acquisition. In Larkin, J. H. and Chabay, R. W., editors, *Computer-Assisted Instruction and Intelligent Tutoring Systems: Shared Goals and Complementary Approaches*, pages 73–109. Lawrence Erlbaum.

Cronbach, L. J. and Snow, R. E. (1977). *Aptitudes and Instructional Methods: A Handbook for Research on Interactions*. Irvington, New York.

del Soldato, T. (1994). Motivation in tutoring systems. Technical Report CSRP 303, School of Cognitive and Computing Sciences, University of Sussex.

del Soldato, T. and du Boulay, B. (1995). Implementation of motivational tactics in tutoring systems. *Journal of Artificial Intelligence in Education*, 6(4):337–378.

Douglas, S. A. (1991). Tutoring as interaction: Detecting and repairing tutoring failures. In Goodyear, P., editor, *Teaching Knowledge and Intelligent Tutoring*, pages 123–147. Ablex Publishing Corporation, Norwood, New Jersey.

Dreyfus, H. (1979). *What Computers Can't Do: The Limits of Artificial Intelligence*. Harper and Row, New York, second edition.

du Boulay, B. (2000). Can we learn from ITSs? In Gauthier, G., Frasson, C., and VanLehn, K., editors, *Intelligent Tutoring Systems: Proceedings of 5th International Conference, ITS 2000, Montreal*, number 1839 in Lectures Notes in Computer Science, pages 9–17. Springer-Verlag.

du Boulay, B., Luckin, R., and del Soldato, T. (1999). The plausibility problem: Human teaching tactics in the 'hands' of a machine. In Lajoie, S. P. and Vivet, M., editors, *Artificial Intelligence in Education: Open Learning Environments: New Computational Technologies to Support Learning, Exploration and Collaboration. Proceedings of the International Conference of the AI-ED Society on Artificial Intelligence and Education, Le Mans France*, pages 225–232. IOS Press.

Eisenberger, R. (1992). Learned industriousness. *Psychological Review*, 99(2):248–267.

Fox, B. A. (1991). Cognitive and interactional aspects of correction in tutoring. In Goodyear, P., editor, *Teaching Knowledge and Intelligent Tutoring*, pages 149–172. Ablex Publishing Corporation, Norwood, New Jersey.

Gegg-Harrison, T. S. (1992). Adapting instruction to the students capabilities. *Journal of Artificial Intelligence in Education*, 3(2):169–181.

Graesser, A. C., Person, N., Harter, D., and the Tutoring Research Group (2000). Teaching tactics in autotutor. In *Modelling Human Teaching Tactics and Strategies: Workshop W1 at ITS'2000, Montreal*.

Grandbastien, M. (1999). Teaching expertise is at the core of ITS research. *International Journal of Artificial Intelligence in Education*, 10(3–4):335–349.

Johnson, W. L., Rickel, J. W., and Lester, J. C. (2000). Animated pedagogical agents: Face-to-face interaction in interactive learning environments. *International Journal of Artificial Intelligence in Education*, 11(1):47–78.

Johnson, W. L. and Soloway, E. (1987). Proust: an automatic debugger for pascal programs. In Kearsley, G. P., editor, *Artificial Intelligence & Instruction: applications and methods*. Addison-Wesley Publishing, Reading, Massachusetts.

Katz, S., Lesgold, A., Eggan, G., and Gordin, M. (1992). Modelling the student in Sherlock II. *Journal of Artificial Intelligence in Education*, 3(4):495–518.

Keller, J. M. (1983). Motivational design of instruction. In Reigeluth, C. M., editor, *Instructional-design Theories and Models: An Overview of their Current Status*. Lawrence Erlbaum.

Koedinger, K. R., Anderson, J. R., Hadley, W. H., and Mark, M. A. (1997). Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence in Education*, 8(1):30–43.

Koedinger, K. R., Corbett, A. T., Ritter, S., and Shapiro, L. J. (2000). Carnegie learning's cognitive tutor: Summary research results. Summary of research results available from CARNEGIE**learning**, Pittsburgh, `www.carnegielearning.com`.

Lajoie, S. P., Wiseman, J., and Faremo, S. (2000). Tutoring strategies for effective instruction in internal medicine. In *Modelling Human Teaching Tactics and Strategies: Workshop W1 at ITS'2000, Montreal*.

Leinhardt, G. and Greeno, J. G. (1991). The cognitive skill of teaching. In Goodyear, P., editor, *Teaching Knowledge and Intelligent Tutoring*, pages 233–268. Ablex Publishing Corporation, Norwood, New Jersey.

Lepper, M., Aspinwall, L., Mumme, D., and Chabay, R. (1991). Self-perception and social perception processes in tutoring: Subtle social control strategies of expert tutors. In Olson, J. and Zanna, M., editors, *Self Inference Processes: The Sixth Ontario Symposium in Social Psychology*, pages 217–237. Lawrence Erlbaum Associates, Hillsdale, New Jersey.

Lepper, M. R. and Chabay, R. (1988). Socializing the intelligent tutor: Bringing empathy to computer tutors. In Mandl, H. and Lesgold, A., editors, *Learning Issues for Intelligent Tutoring Systems*, pages 242–257. Springer-Verlag, New York.

Lepper, M. R., Woolverton, M., Mumme, D. L., and Gurtner, J.-L. (1993). Motivational techniques of expert human tutors: Lessons for the design of computer-based tutors. In Lajoie, S. P. and Derry, S. J., editors, *Computers as Cognitive Tools*, pages 75–105. Lawrence Erlbaum, Hillsdale, New Jersey.

Lesgold, A., Lajoie, S., Bunzo, M., and Eggan, G. (1992). Sherlock: A coached practice environment for an electronics troubleshooting job. In Larkin, J. H. and Chabay, R. W., editors, *Computer-Assisted Instruction and Intelligent Tutoring Systems*, pages 289–317. Lawrence Erlbaum Associates, Hillsdale, New Jersey.

Lester, J. C., Converse, S. A., Stone, B. A., Kahler, S. A., and Barlow, S. T. (1997). Animated pedagogical agents and problem-solving effectiveness: A large-scale empirical evaluation. In du Boulay, B. and Mizoguchi, R., editors, *Artificial Intelligence in Education: Knowledge and Media in Learning Systems. Proceedings of the AI-ED 97 World Conference on Artificial Intelligence in Education*, volume 50 of *Frontiers in Artificial Intelligence*, pages 23–30, Kobe, Japan. IOS Press.

Lester, J. C., Towns, S. G., and Fitzgerald, P. J. (1999). Achieving affective impact: Visual emotive communication in lifelike pedagogical agents. *International Journal of Artificial Intelligence in Education*, 10(3–4):278–291.

Littman, D., Pinto, J., and Soloway, E. (1990). The knowledge required for tutorial planning: An empirical analysis. *Interactive Learning Environments*, 1(2):124–151.

Luckin, R. (1998). 'ECOLAB': Explorations in the Zone of Proximal Development. Technical Report CSRP 386, School of Cognitive and Computing Sciences, University of Sussex.

Luckin, R. and du Boulay, B. (1999). Ecolab: The development and evaluation of a Vygotskian design framework. *International Journal of Artificial Intelligence in Education*, 10(2):198–220.

Major, N., Ainsworth, S., and Wood, D. (1997). REDEEM: Exploiting symbiosis between psychology and authoring environments. *International Journal of Artificial Intelligence in Education*, 8(3–4):317–340.

Malone, T. and Lepper, M. R. (1987). Making learning fun. In Snow, R. and Farr, M., editors, *Aptitude, Learning and Instruction: Conative and Affective Process Analyses*. Lawrence Erlbaum.

McArthur, D., Stasz, C., and Zmuidzinas, M. (1990). Tutoring techniques in algebra. *Cognition and Instruction*, 7:197–244.

McCalla, G. (2000). Life and learning in the electronic village: The importance of localization for the design of environments to support learning. In *Intelligent Tutoring Systems: 5th International Conference, ITS2000, Montreal*, number 1839 in Lecture Notes in Computer Science, pages 31–32. Springer, Berlin.

Merrill, D. C., Reiser, B. J., Beekelaar, R., and Hamid, A. (1992a). Making processes visible: Scaffolding learning with reasoning-congruent representations. In Frasson, C., Gauthier, G., and McCalla, G. I., editors, *Intelligent Tutoring Systems: Second International Conference, ITS'92, Montreal*, pages 103–110, New York. Springer Verlag.

Merrill, D. C., Reiser, B. J., Ranney, M., and Trafton, J. G. (1992b). Effective tutoring techniques: A comparison of human tutors and intelligent tutoring systems. *The Journal Of The Learning Sciences*, 2(3):277–305.

Murray, T. (1999). Authoring intelligent tutoring systems: An analysis of the state of the art. *International Journal of Artificial Intelligence in Education*, 10(1):98–129.

Ohlsson, S. (1987). Some principles of intelligent tutoring. In Lawler, R. W. and Yazdani, M., editors, *Learning Environments and Tutoring Systems: Learning Environments and Tutoring Systems*, volume 1, pages 203–237. Ablex Publishing, Norwood, New Jersey.

Pask, G. and Scott, B. (1975). CASTE: A system for exhibiting learning strategies and regulating uncertainties. *International Journal of Man-Machine Studies*, 5:17–52.

Peachey, D. and McCalla, G. (1986). Using planning techniques in intelligent tutoring systems. *International Journal of Man-machine Studies*, 24:77–98.

Reiser, B. J., Kimberg, D. Y., Lovett, M. C., and Ranney, M. (1992). Knowledge representation and explanation in GIL, an intelligent tutor for programming. In Larkin, J. H. and Chabay, R. W., editors, *Computer-Assisted Instruction and Intelligent Tutoring Systems*, pages 112–149. Lawrence Erlbaum Associates, Hillsdale, New Jersey.

Ridgway, J. (1988). Of course ICAI is impossible ... worse though, it might be seditious. In Self, J., editor, *Artificial Intelligence and Human Learning*, pages 28–48. Chapman and Hall Computing, London.

Rutter, M., Maughan, B., Mortimore, P., and Ouston, J. (1979). *Fifteen Thousand Hours: Secondary Schools and Their Effects on Children*. Paul Chapman Publishing.

Schoenfeld, A. H. (1998). Towards a theory of teaching-in-context. *Issues in Education*, 4(1):1–94. `http://www-gre.berkeley.edu/faculty/aschoenfeld`.

Self, J. (1993). Special issue on evaluation. *Journal of Artificial Intelligence in Education*, 4(2/3).

Shute, V. J. (1995). SMART: Student modelling approach for responsive tutoring. *User Modelling and User-Adapted Interaction*, 5(1):1–44.

Sleeman, D., Kelly, A., Martinak, R., Ward, R., and Moore, J. (1989). Studies of diagnosis and remediation with high school algebra students. *Cognitive Science*, 13(4):551–568.

Soller, A. L. (2001). Supporting social interaction in an intelligent collaborative learning system. *International Journal of Artificial Intelligence in Education*, 12(1):40–62.

Sternberg, R. J. and Horvath, J. A. (1995). A prototype view of expert teaching. *Educational Researcher*, 24(6):9–17.

VanLehn, K. (1987). Learning one subprocedure per lesson. *Artificial Intelligence*, 31(1):1–40.

VanLehn, K., Ohlsson, S., and Nason, R. (1994). Applications of simulated students. *Journal of Artificial Intelligence in Education*, 5(2):135–175.

von Glasersfeld, E. (1984). An introduction to radical constructivism. In Watzlawick, P., editor, *The Invented Reality*, pages 17–40. W.W. Norton & Company, New York.

von Glasersfeld, E. (1987). Learning as a constructive activity. In Janvier, C., editor, *Problems of representation in the teaching and learning of mathematics*, pages 3–17. Lawrence Erlbaum Associates, New Jersey.

Vygotsky, L. S. (1978). *Mind in Society: the Development of Higher Psychological Processes*. Harvard University press, Cambridge, MA.

Vygotsky, L. S. (1986). *Thought and Language*. The M.I.T. Press, Cambridge, MA.

Wasson, (Brecht), B. (1990). *Determining the Focus of Instruction: Content Planning for Intelligent Tutoring Systems*. PhD thesis, Department of Computational Science, University of Saskatchewan, Canada.

Wertsch, J. (1984). The zone of proximal development: Some conceptual issues. In Rogoff, B. and Wertsch, J., editors, *Children's Learning in the "Zone of Proximal Development"*, volume 23, pages 7–19. Jossey-Bass, San Francisco.

Wilson, B. (1997). The postmodern paradigm. In Dills, C. and Romiszowski, A., editors, *Instructional development paradigms*. Educational Technology Publications, Englewood Cliffs NJ.

Wilson, B. and Cole, P. (1991). A review of cognitive teaching models. *Educational Technology Research and Development*, 39(4):47–64.

Winne, P. H. (1997). Experimenting to bootstrap self-regulated learning. *Journal of Educational Psychology*, 89(3):397–410.

Wood, D. J., Bruner, J. S., and Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, 17:89–100.

Wood, D. J. and Middleton, D. J. (1975). A study of assisted problem solving. *British Journal of Psychology*, 66:181–191.

Wood, D. J., Wood, H. A., and Middleton, D. (1978). An experimental evaluation of four face-to-face teaching strategies. *International Journal of Behavioural Development*, 1:131–147.

Woolf, B. P. (1988). Representing complex knowledge in an intelligent machine tutor. In Self, J., editor, *Artificial Intelligence and Human Learning*, pages 3–27. Chapman and Hall Computing, London.