SOME CHALLENGES FOR INTELUGENT TUTORING SYSTEMS*

D Sleeman

Department of Computing Science, The University, Aberdeen, AB9 2UB, Scotland, UK

In this paper, I will outline several critical issues which will have to be addressed before ITSs (Intelligent Tutoring Systems) can become truly responsive to the problems experienced by individual students. Firstly, such systems need the ability to analyse student responses and infer the nature of previously unencountered behaviour patterns. Secondly, even if one has an accurate and highly detailed student model, ITSs need to communicate the nature of the student's errors to the individual student. This turns out to be a difficult problem. These issues will be discussed in detail in my presentation; an outline is given below:

Explaining Previously Unencountered Behaviour

One of the contributions of recent work by workers in the ITS and Cognitive Science fields has been to show that student errors that were said by teachers to be the result of careless work are often due to slighdy buggy procedures.

Brown & Burton (1978) report such bugs for subtraction:

72	123	
28	78	etc
56	155	

where the smaller number is always subtracted from the larger irrespective of its position in the column. Similarly, Sleeman (Sleeman, 1982a) has reported a number of common errors in algebra:

3 + 4x = 9 => x = 9 - 7
5 + 7 x = 24 => x = 24 - 12
(i.e., m + n x = p => x = p - m - n)

2 + 4x = 9 => 2 * 4 + x = 9
4 + 7x = 24 => 4 * 7 + x = 24
(i.e., m + nx = p) => m * n + x = p)

Builders of ITSs have encoded such errors, thus making their systems capable of diagnosing them, (Brown & Burton, 1978; Sleeman, 1982a). Similarly, PROUST is able to analyse deviant student programs by having variants of program plans explicitly represented (Johnson, 1986). The great challenge is to build systems that are able to identify previously unencountered rules. Langley and Ohlsson (Langley & Ohlsson, 1984) address the issue of the student using a known operator inappropriately. This research group has addressed the issue of creating new operators, or at least operators which are variations of known operators. We have evolved two quite distinct approaches:

- The INFER* system attempts to work forward from the initial task, backwards from the student's answer, and infers new rule(s) to fill the "gaps" created by this bi-directional search. (Note the backward search uses a series of heuristics to focus the search.)

The MALGEN system takes a correct rule set, applies a series of perturbations to the rule-set and "filters" the resulting rules to remove those which are thought to be totally unreasonable. (For example, algebra students appear to know the form of an acceptable answer and so we have never seen m x = n transformed to x = n/ or x = An.)

Below we give a trace of the INFER* program as it attempts to infer a new mal-rule for the task 3 + 4 x = 9, given the student answer x = 2:

A: x = 2

B: x = rn+2]-[n]
i

C: x = 9 - 7

D: x + 7 = 9

E: x + [7-n] +[n] = 9

F: x + 3 + 4 = 9

G: 3 + 4 x = 9

Comment on steps

A to B: This is a simple algebraic rearrangement.

B to C: This uses knowledge of the target equation (3 + 4 x = 9) to instantiate n to be a number on the right hand side of the target equation. (This is a heuristic.)

C to D: This uses a domain rule in the backward direction. [Note there are 2 ways this can fire.]

D to E: This step is analogous to the A to B step.

E to F: This step is analogous to the B to C step, but here uses knowledge of the coefficient on the left hand side of the equation to instantiate the variable.

G to F: The *Inference* Step is applied as none of the heuristics or domain rules are applicable.

The above trace shows a successful path. Often in trying to find a new rule in a protocol, the algorithm generates a large number of paths (as of course there are a sizeable number of heuristics). An initial form of this work was described in Sleeman (1982b); a more sophisticated algorithm is discussed in Sleeman, Hirsh & Kim (1987). This current algorithm is able to determine complex numerical relationships between a series of coefficients, and is able to progressively resolve rules once an additional new rule has been encountered.

E1XIE and Model-based Remediation

It has long been assumed by workers in ITS that once one has a complete model of a student's problem solving on a range of tasks, then remediation of the student would be straightforward, (Brown & Burton 1978). Recent experiments with the PIXIE project have attempted to demonstrate this.

As a result of earlier analyses of students' errors in solving algebra tasks we were able to formulate a number of errors or mal-rulcs. These, together with correct rules, and a series of tasks formed a database for the PIXIE Intelligent Tutoring System (Sleeman, in press). PIXIE is able to infer a model of the student's problem solving expressed in terms of these rules and mal-rules. In our earlier experiments, we merely provided these analyses to the teacher. Subsequently, we watched a teacher remediate individual students based on these analyses, and after interviewing a number of other teachers, implemented a remedial system based on this experience. (Virtually all the teachers we encountered said that they taught algebra in a procedural, or rule-based, way.)

The form of the remediation now provided by PIXIE is:

highlight each of the incorrect transitions in the student's protocol (e.g., "you appeared to change 3x+5=9 to x+3+5=9"), explain why that was wrong and then say what the change should have been. (This step is repeated for each error in the protocol.)

show the student with an appropriate commentary how he should have solved the complete task.

When this was first tried out with a small number of students, we were told:

the comments were too verbose,

certain mathematical terms used in our explanations were not understood. [We have now provided a flash card.]

there were difficulties because PIXIE's standard approach was to ask the student to work a series of tasks, then, if necessary to start remediating from the first task; by which time the student had forgotten how he had worked the first task. [We now remediate, immediately after an error, when necessary.]

These changes were made to PIXIE and we ran an experiment to determine whether simply RETEACHING a topic was as effective as TARGETED (or MODEL-BASED REMEDIATION) ie. the remedial sequence discussed above. As this result was inconclusive we decided that we should attempt a series of experiments using human-tutors to probe these phenomena. A summary of these experiments is:

no difference has been found between the effectiveness of RETEACHING and TARGETED remediation - even when we modified the TARGETED approach to include Cognitive Dissonance (i.e., demonstrating to the pupil his solution was wrong) and Cognitive Engagement (involving the student actively in explaining what he had done).

the errors with this population appear to be unstable. That is errors, noted on a pretest often have disappeared by the time of the interview several days later.

Summary

Getting a system to recognise previously unencountered behaviour patterns, is a very open-ended problem which is, in general, not-solvable.

Even though PIXIE has a model for a student's problem solving, it has not so far proved possible to remediate very effectively. There appears to be two reasons for this. Firstly, the errors shown by this student population on this topic appear to be unstable. (Thus an important precondition for remediation is invalidated.) Secondly, we now believe that remediation is very complex - even when one knows what one wishes to communicate. After all, much of our social know-how is deciding how to say something to another person so we can be understood, and secondly so that the listener will do what the speaker wishes him to do. (Frequently, we have to go to extreme lengths to achieve this. Such as asking a known competitor to do the task, thereby the person himself wishes to do the task.) Cronbach and Snow (1977) have already demonstrated the difficulty of deciding an effective treatment for a student on the basis of a student's aptitudes. Their work was based on post-hoc analyses of the experiments of a series of investigators; thus ITS workers had some justification for believing those conclusions were unduly pessimistic. However, the recent experiments with PIXIE suggest that their conclusions are probably worryingly realistic!

REFERENCES

Brown, J. S. *Sc* Burton, R. R. (1978). Diagnostic models for procedural bugs in basic mathematical skills. Cognitive Science. 2* pp 155-192.

Cronbach, L J & Snow, R E (1977). Aptitudes and Instructional Methods. New York: Irvington.

Johnson, W L (1986). Intension-based Diagnosis of. NOVICE Programming ERRORS London: Pitman (Research Notes in AI scries).

Langley, P & Ohlsson, S (1984). Automated Cognitive Modelling. proceedings of the AAAI *z.* 84 Conference, PP 193-197.

Sleeman, D (1982a) Assessing aspects of competence in basic algebra. In Intelligent Tutoring Systems (eds D Sleeman & J S Brown) London: Academic Press, pp 185 - 199.

Sleeman, D (1982b). Inferring (mal)rules from pupils* protocols. In the Proceedings of the 1282 European AI Conference, pp 160-164.

Sleeman, D (in press). PIXIE: A shell for developing Intelligent Tutoring Systems. In Intelligent cAI (M. Yazdani & R. Lawler, eds). New Jersey: ABLEX.

Sleeman, D, Hirsh, H B & In Yung, Kim (in prep). Expanding an Incomplete Domain Theory: Two Case Studies.

ACKNOWLEDGEMENTS

Support for this work has been provided by IBM and by ARI/ONR under contract number MDA 903-84-K-0279. The recent experimental studies have been carried out with Rosemary Martinak, Eamonn Kelly, and Joyce Moore.

* Unfortunately, only an extended abstract for my invited talk is available at the time of going to press. A full paper will be available from the author subsequently.