By Professor
Sally Yeates Sedelow '53BA

Computers and

"So what's wrong with the way I write?" When the student asks this question, the English teacher points to misspelled words, or a plural noun and singular verb, or a modifying phrase which modifies nothing. But if the spelling and grammar are, in fact, correct, what then? Well, perhaps the writing is monotonous: every sentence begins with "The," or begins with a subject followed by a verb, or has about the same number of words.

But suppose there is sufficient variation to relieve monotony — what then? Then the teacher points to the ideas chosen and their development; the student responds that Nobel prize winners have written about the same ideas; the teacher says, yes, but the Nobel laureates did a better job and the student wants to know what "better" means. Perhaps the teacher can find an answer that satisfies the student but often, at this stage of a discussion, neither student nor teacher is content with the answer. Both are struggling with general value-laden words — better, truer; or words implying something about perception — more accurate, more realistic; or about performance — more mature, more calculated. As is obvious, the meaning of such words is very difficult to define. So that it is hard for the teacher to convey his meaning to the student, indeed, it is difficult for the teacher to specify to himself the reasons for his feeling that the student could do a better job.

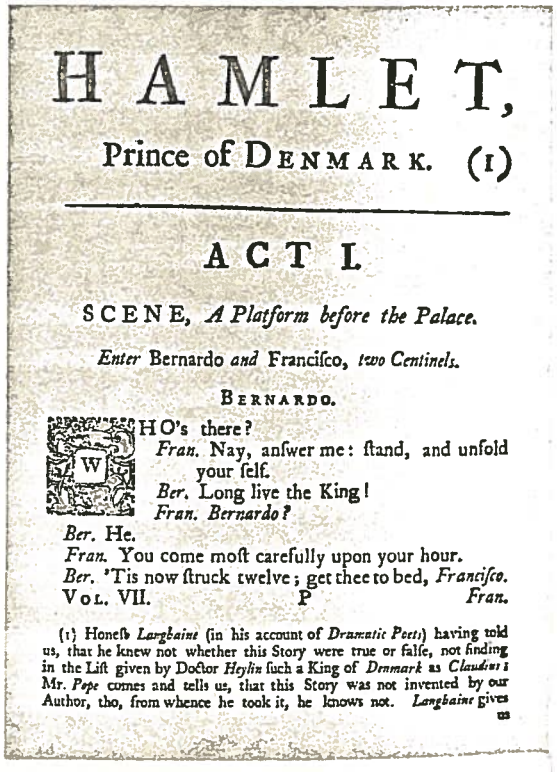
Teachers of literature frequently find themselves in the same box. The teacher likes a poem by Milton and the student does not; presumably the teacher is more agile at specifying what is good than the student at specifying what is bad. The student is persuaded — until he encounters another teacher who tells him how bad that particular poem by Milton is.

What then? Confusion for the student. Criteria for evaluation are so vague that two presumed experts can arrive at opposite conclusions. If the student is asked to write a paper on that poem, what will he say?

Such questions are prevalent throughout those disciplines known as the humanities, however well-buried the questions may be beneath the elegant, or straightforward, or abstruse, or hard-hitting, prose of practitioners in these disciplines. Such questions have, in part, prompted an interest in the use of the computer by a small but growing number of scholars in the humanities. Why the computer? Can a machine that "thinks" in terms like these — 010 001 — tell us anything about Shakespeare? Or can it help teach a student to write?

First of all, we shouldn't boggle at the 01 symbol-system. Our ABC symbol-system doesn't, *per se*, look much like the writings of Shakespeare either. It's the language built from the symbol system that matters. And since the designers of the computer's efficient 01 symbol-system noticed the importance of the ABC system to much data the computer might study or generate, the ABC system can be translated into the 01 system and vice versa. For example, 010 001 is one computer's translation, or representation, of the letter A; 010 010 of the letter B; and so on. Clearly, the writings of Shakespeare can be translated into a form so that the computer can "study" them.

Granted, then, that the computer can produce a copy of Shakespeare's writings for its own study; what kind of student is the computer? Answers to this question range from glowing de-



HAMLET,

Prince of DENMARK. (1)

ACT I.

SCENE, *A Platform before the Palace.*

Enter Bernardo *and* Francisco, *two Centinels.*

BERNARDO.

WHO's there?

Fran. Nay, answer me: ſtand, and unfold your ſelf.

Ber. Long live the King!

Fran. Bernardo?

Ber. He.

Fran. You come moſt carefully upon your hour.

Ber. 'Tis now ſtruck twelve; get thee to bed, *Francisco.*

(1) Honeſt *Langbaine* (in his account of *Dramatic Poets*) having told us, that he knew not whether this Story were true or falſe, not finding in the Liſt given by Doctor *Heylin* ſuch a King of *Denmark* as *Claudius*: Mr. *Pope* comes and tells us, that this Story was not invented by our Author, tho, from whence he took it, he knows not. *Langbaine* gives us

Language

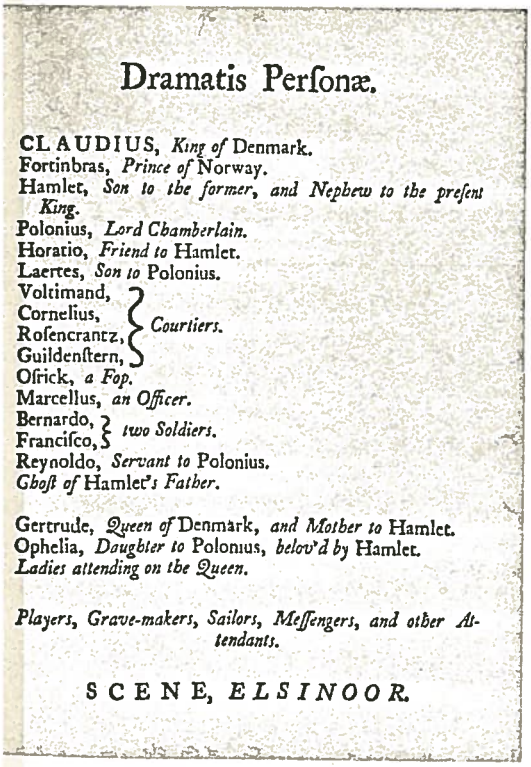
scriptions of the computer's ability to defeat champion checker players to bored indications of its penchant for doing the same thing again and again and again until the money to pay it for doing its dull, repetitive task runs out or until the task is concluded. To an extent — and more than a human student — a computer is as smart or as dumb as we either want it to be or know how to tell it to be.

Humans, made nervous by the computer, often say with relief: "But a computer only does what you tell it to do." True. And it's the type and degree of specification, taken together with the "experience" or data given the computer, which determine its mode of behavior. For example, the computer can be instructed to count the number of "the"s in a two million word text; for this task, it needs to have a "the" in its memory (so that it can compare each word in the text with "the"); it needs to have a copy of the text; and it needs instructions telling it how to pick out each individual word in the text. Without the latter, it might think segments like ter, ter i, text., or ithou were individual words.

To play checkers, the computer must have a good deal more information and it must keep applying that information to a changing situation. As the positions of the checkers change on the board, new rules apply or rules already used must be utilized for a new situation. Sometimes the computer is presented with a situation, such as a game of checkers, and a series of rules which might be used to play such a game. When the computer selects a rule (or sequence of instructions) correctly, that behavior is positively reinforced, so that in a similar situation that rule is likely to be used again. On the other hand, if the computer chooses badly, negative reinforcement guarantees that the rule is less likely to be chosen the next time a similar situation occurs.

All of the examples above represent approaches — which human beings use and know how to specify — to performing a task. However, in the first example, each action of the computer can be exactly predicted.

(continued on page 18)



Dramatis Perſonæ.

CLAUDIUS, *King of* Denmark.
Fortinbras, *Prince of* Norway.
Hamlet, *Son to the former, and Nephew to the preſent King.*
Polonius, *Lord Chamberlain.*
Horatio, *Friend to* Hamlet.
Laertes, *Son to* Polonius.
Voltimand, Cornelius, Roſencrantz, Guildenſtern, *Courtiers.*
Oſrick, *a Fop.*
Marcellus, *an Officer.*
Bernardo, Francisco, *two Soldiers.*
Reynoldo, *Servant to* Polonius.
Ghoſt of Hamlet's *Father.*

Gertrude, *Queen of* Denmark, *and Mother to* Hamlet.
Ophelia, *Daughter to* Polonius, *belov'd by* Hamlet.
Ladies attending on the Queen.

Players, Grave-makers, Sailors, Meſſengers, and other Attendants.

SCENE, *ELSINOOR.*

ABOUT THE AUTHOR

Sally Yeates Sedelow is Professor of Computer Science and Linguistics at the University of Kansas in Lawrence. A 1953 alumna of The University of Iowa, she took her MA at Mount Holyoke College in 1956 and has a 1960 PhD from Bryn Mawr.

Professor Sedelow is a member of an international group of Shakespeare scholars who are holding The World Shakespeare Congress in Vancouver, British Columbia, Canada, this summer. She has served on numerous panels and editorial boards in the field of computer teaching and research, and is a field reader for the U.S. Department of Health, Education and Welfare, and for The Canada Council.

During 1970-71 Professor Sedelow was a Visiting Scientist for the Association for Computing Machinery of the National Science Foundation, lecturing widely on computer use in humanities during that time. A native of Greenfield, Iowa, she was married in 1958 to Walter A. Sedelow, Jr., who now works with her at the University of Kansas.

(continued from page 7)

It is simply going to compare each word in the text with its copy of "the" until it has looked at every word in the text. Whenever the word in the text matches "the," the computer will add 1 to the number indicating how many "the"s it has already counted. The only unknown is just how many "the"s there will be.

In the checker game, each move by the computer cannot be predicted because the computer will be responding to the other player's moves. Further, it may be programmed to respond in any one of a number of ways in any given situation. So, in the checker game, the computer behaves according to options which are precisely specified by its human programmer — but it is not possible to predict which option it will choose.

To recur to the earlier question — what kind of student is the computer? — we can safely say that the computer is a more diligent counter of "the"s than the average student is likely to be, and we can also say that its experience as a checker player has been so enriched by detailed specifications of options by good human checker players that it is a better checker player than most human beings. Not only does it have good instructions, but it can bear them all in mind as needed.

What is the relevance of speculation about the computer's talent as a student, for improving the teaching of writing or of literature? Namely, that the computer can help give precise meaning to vague terms by: 1) totting up information about language behavior which would be too tiresome for a human being to accumulate; 2) keeping in mind and acting upon a great range of instructions or rules for action, including the details associated with each instruction or rule. Thus the computer can serve as one type of research assistant for the faculty member.

To take a case in point, the computer could be given an essay by the student and by a Nobel laureate on the same topic and it could begin by accumulating a great many "boring" statistics about all elements which can be described statistically. Averages for word length, sentence length, and paragraph length could be computed for both the student and Nobel laureate. Sequences of events (e.g., five sentences with twelve words each, followed by five with six words, etc.) could also be tabulated, as could aspects of sequences. The latter would include such possibilities as every sentence beginning with "the" or every sentence beginning with a subject followed by a predicate.

To identify categories such as subject or predicate, the computer would either have to be told in each instance, or it would have to produce syntactic parsings of the text. Many attempts have been made to specify rules which would enable the computer to produce such parsings. Some of these attempts have been reasonably successful, but no parser exists that can deliver 100% accuracy. But then, human parsers aren't perfect either.

In summary, the computer can be used to look for a range of characteristics of writing that might form the substance for more vague evaluative words such as monotonous, or repetitive, or immature. Furthermore, it can look at such characteristics so thoroughly that many which might escape notice by the human eye are revealed and can thus be relayed to the student (and his teacher).

Exploring the ideas present in a text, and their development, is a task for which the specification would greatly exceed that of a checker game — for the wonderfully vague realm of semantics is involved. Too little is known about semantic relationships among words to be able to tell a computer exactly how to look for ideas. Nonetheless, approaches do exist which have produced rather good results.

VIA, a system of computer programs which forms part of my own research efforts, is just such an approach. VIA stands for Verbally-Indexed Associations and is designed to look for themes, or central motifs, or central concepts or ideas in any text. To operate, VIA must have a procedure for grouping words together by root (e.g., putting together "book" "books" "bookish"), and then for grouping them according to wider semantic relationships. For the latter, the computer must either have an enormous memory with lightning retrieval capacity (like the human brain), or it must be given some kind of summary of word usage which makes such relationships apparent.

A thesaurus is that kind of summary, and VIA is given such a thesaurus in order to link words together on a semantic basis. VIA is designed so as to pull together extensive groups of semantically-related words. In fact, no human could deal comfortably, if at all, with so much detail. Put to the test on Shakespeare's *Hamlet*, VIA turned up most if not all the themes literary critics have commented on (e.g., the many words connected with disease and decay, as in "Something is rotten in the State of Denmark"), as well as others which seem to contribute to an interpretation of the play, but which no one had noted.

In complexity, VIA is more analogous to the checkers game than to counting "the"s. Because — although VIA's operations are specified — it is, practically speaking, impossible to predict which words in the text will be pulled together into semantic clusters. The corpus of a language is too great for the human being to bear in mind all the individual words.

At present the research effort on which my husband, Walter, and I, and a number of graduate students, are at work is directed toward combining the programs designed for studying the semantics of a text with programs for exploring those elements of the text which can be described sta-

tistically. Ideally, this combination or "package" of programs would be very like the checkers game. That is, the computer would choose a program or sequence of programs on the basis of results from a program that had already operated. Thus the options would be specified but the choice of option would depend upon the given situation.

If we are able to design such a system, it will provide powerful assistance to the teacher who wants to "flesh out" vague generalizations about writing or literature with specific detail.

In time, too, we hope to provide computer systems which will complement the teacher in other ways. For example, it is very dificult to talk about the auditory component of literature; both students and teacher yawn if patterns of alliteration or of internal rhyme are pointed out at length with any detail.

Yet the sound of literature is extremely important. Why not design computer programs which will provide representations of auditory elements, so that the student can request a graph or display of the behavior of certain sounds in a given poem? Many of these "assists" will someday be possible.

Because of its brevity, this article has ignored the practical and theoretical problems which must be solved in order to facilitate easy interaction between computer, teacher, and the student who wants to improve his writing or his analysis of literature. Nonetheless, the potential and value of the computer for this presently hazy pedagogical area should be clear. Steps which have already been taken make ultimate fulfillment of its promise seem worth working for.

The U of I Art Museum shows "Accessions 1970-1971," 130 newly-acquired art works, until June 27. Late summer showings will concentrate on permanent Museum possessions. Hours, 10:30-5 Mon.-Fri.; 10-5 Sat.; 1-5 Sun. Free.