STROLLING DOWN THE GARDEN PATH: ERROR PRONE TASKS IN EXPERT PROBLEM SOLVING

Paul E. Johnson
William . Thompson

University of Minnesota

ABSTRACT

Problems which induce performance that has the false appearance of success (garden path problems) may be an inevitable consequence of the need (in both human and computer problem solving systems) to create abstract knowledge representations in order to make problem solving efficient. An example is presented from s domain of physics problem solving tasks in which a hierarchical organisation of lines of reasoning leads to errors of the garden path type. Several aspects of a possible model of the problem solving process in these tasks are briefly outlined.

1. INTRODUCTION

Classes of problem solving tasks exist in which errors regularly occur, even among (human) expert problem solvers. The usually desirable strategy of beginning a task with fairly global concepts and then working towards more specific levels of detail (1] frequently fails when the structure of the problem is such that the initial choice of global representation either does not apply or else fails to emphasise the detail of knowledge required for a successful solution. Some types of problem structures have the additional property that the dominant global strategy they elicit appears to be correct when, in fact, it is not. Problems of this type can be refered to as "garden path problems** since they induce performance that has the false appearance of success.

If garden path problems consisted merely of pussies or tasks in which tricks or caveats had to be discovered in order for a correct solution to be obtained, they might be interesting curiosities, but the attention given to them by the scientific community would be small. What makes these problems interesting, however, is the fact that they may exist in a wide variety of knowledge domains (eg. (2]). Thus, they are of relevance to the study of both cognitive science and expert systems.

This work was supported in part by the National Science Foundation under Grant SED79-13036.

2. AN EXAMPLE: MILKO

One famous example of a garden path problem is found in the domain of physics problem solving and is attributed to Harvey Cohen [3]* In Cohen's original problem (termed MILKO), a milk bottle is filled with a mixture of milk and cream and left to stand. After a time the milk and cream separate and the cream rises to the top (without any change in volume). Cohen asks simply whether the pressure at the base of the bottle is the same or different after separation of the two fluids. Expert physicists, as well as more novice problem solvers, frequently fail to correctly solve the MILKO problem.

Difficulties with the MILKO problem appear to be due primarily to improper choices of problem representation and level of detail in reasoning. Pressure change is associated with change in either the height or amount of fluid above the base, and since neither of these change in the MILKO problem, there is presumably no change in the pressure (the garden path answer). What is not considered, however, is that the problem deals with two variables: non-homogeneous fluids and a container with slanted sides (the milk bottle). If one of these variables is present, by itself, there is no pressure change (Pascal's Principle). If both variables are present, however, then the pressure at the base of the container is less following the separation of the two fluids.

Understanding competence in garden path problems has implications for both cognitive modelling and error detection/recovery in formal problem solving systems. For example, if following the MILKO problem, individuals are given a second problem in which shape is made explicit by a drawing of a flask with slanted sides, the more expert individuals will often recognise their error in the first problem and proceed to do both problems correctly. For some experts and most non-experts, however, the illusion that the pressure does not change is so great that a third problem is necessary - one in which explicit values are given for the variables of shape and fluid densities, so that exact calculations can be made.

Detection of errors occurs when individuals in the second problem take explicit notice of the slanted sides of the container, or in the third problem are faced with calculations that show that the pressure In the separated state is less. Recovery from error In either esse consists of choosing sn slternatlve problem representation, a more detailed line of reasoning, or both. When experts working on problems of this type see errors In their thinking, they not only recover end sdopt the correct approach to the lowedlate problea, but they also frequently generalise this recovery both backwards and forwards. Not only Is the previous error recognised, but additional problems of the same type are solved correctly. Novices, by contrsst, do not generalize either backward or forward and will frequently abandon a correct approach when given a new problea.

2' A FRAMEWORK FOR MODELLING

Observstions of human performance on error prone tasks provide important Information for the development of formal models of problem solving. Not only can the limits of competency be investigated, but it is often essier to probe for the underlying mechanisms when the problem solving process hss been perturbed by applying it to situations likely to result in incorrect decisions [4]. Studying the differences in performance between experts snd novices on the same class of tssks gives additional insight by allowing comparisons of problem solving using different knowledge bases. We heve been investigating the performance of human subjects on a variety of error prone problem solving tssks in physics with the sim of developing s computational model of this process. Our preliminary reaults indicate that a hierarchical planning process [5] is employed in solving these problems. Unlike previous work in this area, however, we find that there is no clear level of primitive actions at which the validity *of* a plan can be confirmed.

3.1. Hierarchical lines of reasoning

To solve all but the most trivial problems, Individual actions must be organised In an efficient way. For many teak domains, expert (human) problem solvers often structure their knowledge Into "lines of reasoning" - partially or fully ordered sequences of conceptual actions which are useful for specific types of problems [6,7]. Significant efficiencies arise becsuse these sequences are predetermined and need not be rediscovered for each new variant of a problem. *A* general claas of teaks will typically have many lines of reasoning associated with It. These lines of reasoning can be characterised by the categories of knowledge used and by the level of detail with which that knowledge Is represented. Problem solving occurs by identifying an adequate line of reasoning, mapping the problem specification into the appropriate repreaentatlon, and then executing the implied Inference steps.

If problem solving knowledge is structured In terms of levels of detail, then It is often possible to further reduce the set of actions which must be searched In order to find a solution. Problems can be first solved at an abatract level of representation and thla solution can then uaed to constrain the possibilities at more detailed levels (5J. The garden path errors In our example occur becsuse many problems sre approached using s hierarchy of lines of ressoning n which the more detailed units of knowlsdge sre often never exsnlned. Such en organisation appears necesssry for solving complex problems with limited computational resources. Efficiency is schleved by terminating the sesrch through the hlersrchy whenever it sppesrs thst the (sbstrsct) solution is likely to be correct. Thus, garden path errors represent examples of the failure of powerful heuristics thst, for the most psrt, grestly extend the effectiveness of s problem solving system.

A general class of tasks will typically have many lines of ressonlng sssoclated with it. These lines of reasoning can be characterized by the categories of knowledge used and by the level of detail with which that knowledge is represented. For example, two major lines of reasoning are suggested by the MILKO problem. One is bssed upon pressure, forces, and cross sectional area and is formalised by P - F/A. The other is based upon pressure, density, gravity, and fluid height and is formalised by P - pgh. Either line of reasoning lesds to failure when applied at the global level where concepts in the corresponding equations are instantiated in terms of the total fluid in the container. Either line of reasoning can lead to success when applied at a more detailed level where concepts are specified in terms of decomposed problem states.

3.2. Control Issues

The manner in which control is exercised within snd among lines of ressonlng is centrsl to the modeling of expert problem solving behavior. The desirable breadth of search has been a continuing topic among the artificial Intelligence community. Nsrrowly structured sesrehes promise potential efficiency, but easily result In garden path errors. Brosder search sacrifices "best-case" efficiency for Improvements in accuracy. Our studies of human problem solving suggest that breadth of search is s function of both problem type and level of expertise [8]. 111 structured problems or those for which limited knowledge (expertise) is available will result in less effective heuristics for recognising the appropriate line of reasoning. This must be compenssted for by control functions thst lesd to greater breadth in the seerch process snd thus allow for the simultaneous considerstion of more possibilities.

In human problem solving, differences appear In the ways in which experts snd novices implement lines of ressonlng. If a novice abandons a particular line of reasoning snd then returns to it later in the problea solving process, the line of ressonlng Is usually reinvestigated from the beginning. Both the space ssvlngs snd time inefficiencies of clssslcsl bscktrsck programming sre present. Few Intermedlete results need be remembered, but many sub-problems may- be re-eolved severel different times. Experts, on the other hand, seem to utilise some sort of coroutine structure. Lines of reasoning are reactivated essentially at the point at which they were suspended. This Bay be possible because nore finely developed control heuristics result In consideration of far fewer possibilities, allowing a small set of partial results to be preserved.

4. SUMMARY

Garden path problems have been Identified and studied In task domains as diverse as physics and medicine [8]. Such problems may be a natural consequence of hierarchically organised lines of reasoning created to support expert problem solving. We hope that our analysis can eventually lead to Improved frameworks for knowledge engineering efforts as well as more complete models of cognitive behavior In problem solving.

ACKNOWLEDGMENTS

Robert Sedlmeyer and Richard Keller provided valuable comments during the preparation of this paper.

BIBLIOGRAPHY

[1J J.H. Larkln, "Teaching problem slovlng in physics: The psychological laboratory and the practical classroom,** in Problem Solving In Education, D.T. Tuma and F. Relf, ede., Hillsdale, N.J.: Lawrence Erlbaum Associatea, 1960.

(2) R. Pelerls, Surprises in Theoretical Physics, Princeton: Princeton University Press, 1979.

(31 H.A. Cohen, "The art of snaring dragons," MIT AI Memo 338, 1975.

(4) H.A. Simon, The Sciences of the Artificial, Cambridge: MIT Press, 1961.

[5]E.D. Sacerdotl, A Structure for Plans and Behavior, New York: Elsevier North-Holland, 1977.

(6; E.A. Felgenbaum, "The art of artificial intelligence: Themes and case studies of knowledge engineering,** Proc. Fifth International Joint Conference on Artificial Intelligence. 1977.

(7) J.H. Larkln, J. McDermott, D.P. Simon, and H.A. Simon, "Expert and novice performance In solving physics problems,*' Science, vol. 208, pp. 1335-1342, 1980.

[8] P.E. Johnson, A. Barreto, F. Hassebrock, J. Holler, M. Prletula, P. Feltovlch, and D. Swanson, "Expertise and error In diagnostic reasoning,** Cognitive Science, 1981 (In press).