



Western Michigan University
ScholarWorks at WMU

Dissertations

Graduate College

6-1996

An Account of Expert Phylogenetic Tree Construction from the Problem-Solving Research Tradition in Science Education

Steven D. Brewer
Western Michigan University

Follow this and additional works at: <https://scholarworks.wmich.edu/dissertations>



Part of the Genetics and Genomics Commons

Recommended Citation

Brewer, Steven D., "An Account of Expert Phylogenetic Tree Construction from the Problem-Solving Research Tradition in Science Education" (1996). *Dissertations*. 1697.

<https://scholarworks.wmich.edu/dissertations/1697>

This Dissertation-Open Access is brought to you for free and open access by the Graduate College at ScholarWorks at WMU. It has been accepted for inclusion in Dissertations by an authorized administrator of ScholarWorks at WMU. For more information, please contact wmu-scholarworks@wmich.edu.



**AN ACCOUNT OF EXPERT PHYLOGENETIC TREE CONSTRUCTION FROM
THE PROBLEM-SOLVING RESEARCH TRADITION
IN SCIENCE EDUCATION**

by

Steven D. Brewer

**A Dissertation
Submitted to the
Faculty of The Graduate College
in partial fulfillment of the
requirements for the
Degree of Doctor of Philosophy
Department of Science Studies**

**Western Michigan University
Kalamazoo, Michigan
June 1996**

**AN ACCOUNT OF EXPERT PHYLOGENETIC TREE CONSTRUCTION FROM
THE PROBLEM-SOLVING RESEARCH TRADITION
IN SCIENCE EDUCATION**

Steven D. Brewer, Ph.D.

Western Michigan University, 1996

Content knowledge, strategic knowledge, and forms of knowledge organization used in phylogenetic tree construction are described to account for expert performance. A literature review is used to situate this research with respect to studies of: (a) the teaching of evolution, (b) problem-solving expertise, (c) physics and biology problem solving, and (d) the practice of phylogenetic systematics. Rationales are provided for the development of a computer-based problem-solving environment, model problems, and research problems and methodology.

Each problem consisted of a matrix of coded and polarized phylogenetic data organized by taxa and characters. Four series of research problems were constructed that varied the numbers of solutions, taxa, and characters. In addition, some problems contained a revision component that required: additions to prior solutions, restructuring of prior solutions, or increased or decreased numbers of solutions. Nine expert phylogenetic systematists participated in the research project by thinking aloud while constructing phylogenetic trees to account for the problem data matrices.

A procedural model of expert performance for phylogenetic tree construction is described with associated strategies and heuristics. Three overall strategies were described: inclusion/exclusion, order of divergence, and duplicated taxa. Each overall strategy is represented as a pathway through the model. In addition, a new conceptual

framework for understanding the nature of phylogenetic tree construction problems is described including factors that contribute to difficulty of phylogenetic problems.

The strategies related to the procedural model of expert performance are used to provide an outline of a model of desired performance. The model provides a basis for developing an approach to teaching evolution based on effects-to-causes problems. Examples are provided that illustrate how the results of this research can inform the development of a problem-based approach to teaching evolutionary biology. In addition, the implications for improved problem-solving environments for phylogenetic biology are described.

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

**A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
313/761-4700 800/521-0600**

UMI Number: 9636862

UMI Microform 9636862
Copyright 1996, by UMI Company. All rights reserved.

**This microform edition is protected against unauthorized
copying under Title 17, United States Code.**

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

ACKNOWLEDGMENTS

Three people, more than any others, contributed to the completion of my dissertation. I offer my deepest thanks and sincerest respect for my advisor, Dr. Robert Hafner, who understood at the beginning where this project could lead, but who left it to me to find my own way to follow it. I am both grateful and honored that Dr. Daniel R. Brooks, one of the pioneers of phylogenetic systematics, volunteered such a large measure of his seemingly scarcest resource, time, to serve on my committee and lend me his experience and expertise. Finally, I dedicate my dissertation to my wife Alisa. Words cannot express my gratitude for the love, devotion, and support which she has given me.

Steven D. Brewer

TABLE OF CONTENTS

ACKNOWLEDGMENTS	ii
LIST OF TABLES	vii
LIST OF FIGURES	viii
CHAPTER	
I. INTRODUCTION	1
Teaching of Evolution.....	1
Problem Solving.....	2
Phylogenetic Investigator.....	3
Problem Statement	4
Significance.....	6
II. LITERATURE REVIEW	11
Introduction	11
Toward a Problem-Based Instructional System.....	14
Expertise.....	15
The Study of Expertise.....	15
The Expertise Approach.....	16
The Nature of Expertise	19
A Study of Experts in Genetics	21
Diagrammatic Reasoning in Meiosis	25
A Brief History of Systematics	25
Phylogenetic Inference	27
Principles of Phylogenetic Inference	29
Selection of Ingroup and Outgroup.....	31

Table of Contents—Continued

CHAPTER

Selection of Characters and States	31
Determination of Polarity	32
Phylogenetic Tree Construction	34
Toward a Theory of Expertise in Phylogenetic Analysis	36
III. METHODOLOGY	37
Introduction	37
Analysis of Task	37
The Problem-Solving Environment	38
A Process Model of Evolution	38
Phylogenetic Investigator	40
Rational Analysis	54
Defining the Nature of the Problems	56
A Typology of Model Phylogenetic Problems	57
Research Problems	64
Data Gathering	90
Participants	91
Procedures	93
Data Analysis	97
Development of a Model of Expert Performance	98
Intercoder Reliability	100
IV. RESULTS	102
Introduction	102
Phylogenetic Investigator	103

Table of Contents—Continued

CHAPTER

Strategic Knowledge	106
A Procedural Model of Expert Phylogenetic Tree Construction	107
Components of the Model	109
An Example Problem	133
Strategies of Phylogenetic Tree Construction	163
Content Knowledge	182
Organization of Knowledge	184
Success in Problem Solving	187
Variability Across Participants	187
Variability Across Problem Types	191
Errors	196
Participant S2	203
V. DISCUSSION	206
Toward a Model of Desired Performance	207
Theoretical Significance	209
Problem Solving and Expertise	209
Problem Solving in Biology	214
Practical Significance	222
Phylogenetic Investigator	223
SCI 170	225
Questions for Future Study	228
Conclusion	230

Table of Contents—Continued

APPENDICES

A. A Primer of Phylogenetic Assumptions, Diagrammatic Elements, and Terms.....	231
B. Protocol Clearance From the Human Subjects Institutional Review Board	238
C. Internet Recruitment Posting	240
D. Letter of Recruitment	244
E. Research Checklist	246
F. Letter of Informed Consent	248
BIBLIOGRAPHY	250

LIST OF TABLES

1.	A Comparison of Characteristics of Physics and Genetics Problems and Problem-Solving	23
2.	Methods Necessary for Answering Phylogenetic Questions	28
3.	Objects, States, and Processes That Underlie Phylogenetic Investigator	41
4.	A Rational Analysis of Phylogenetic Tree Construction	55
5.	Summary of Research Problems	67
6.	Problems Attempted (A) and Completed (C) by Participant	96
7.	A Key to Program Operations Used in Example Problem.....	134
8.	Problem 5.3 as Solved by S1 Interpreted Using the Model of Expert Phylogenetic Tree Construction	135
9.	Problem 5.3 as Solved by S7 Using Duplicated Taxa as a Strategy	170
10.	Problem 5.3 as Solved by S8 Considering Order as a Strategy	176
11.	Number of Optimizations per Topology by Problem and Subject.....	188
12.	Average Success in Percent by Participant Across Problems and Topologies.....	191
13.	Data for Problem 1.1 as Redescribed by S2.....	204
14.	Representative General Heuristics With Instantiations in Phylogenetic Tree Construction.....	212
15.	Model-Revising Heuristics With Instantiations in Phylogenetic Tree Construction	217

LIST OF FIGURES

1.	Equivalent Ladderized (Upper) and Non-Ladderized (Lower) Cladistic Representations Created Using MacClade 2.1	47
2.	Cladistic Hypothesis Created Using Phylogenetic Investigator	48
3.	A Phylogenetic Hypothesis Incorporating Fossils as Ancestral (F01) and Sister (F02) Taxa.....	49
4.	A Theoretical Hypothesis Showing R01 as a Living Fossil and an Increased Rate of Change Between PD and PE	51
5.	A Theoretical Hypothesis Showing Punctuated Equilibrium With One Period of Speciation	52
6.	A Theoretical Hypothesis Showing Hybridization	53
7.	Applying Rational Analysis in Problem Solving	56
8.	An Autapomorphy Data Matrix	57
9.	A Phylogenetic Tree Representing the Autapomorphy Problem.....	58
10.	A Synapomorphy 1 Data Matrix	58
11.	A Phylogenetic Tree Representing the Synapomorphy 1 Problem.....	59
12.	A Synapomorphy 2 Data Matrix	59
13.	A Phylogenetic Tree Representing the Synapomorphy 2 Problem.....	60
14.	A Synapomorphy 3 Data Matrix	60
15.	A Phylogenetic Tree Representing the Synapomorphy 3 Problem.....	61
16.	A Homoplasy 1&2 Data Matrix	62
17.	The Convergence Character Optimization for the Homoplasy 1&2 Problem	62
18.	The Reversal Character Optimization for the Homoplasy 1&2 Problem	63
19.	A Homoplasy 3 Data Matrix	63
20.	A Phylogenetic Tree Representing the Homoplasy 3 Problem.....	64
21.	A Homoplasy 4 Data Matrix	64

List of Figures—Continued

22.	Solution 1 of Two Equally Parsimonious Solutions for the Homoplasy 4 Problem	65
23.	Solution 2 of Two Equally Parsimonious Solutions for the Homoplasy 4 Problem	65
24.	The Data Matrix for Problem 1.1	69
25.	The Most Parsimonious Solution for Problem 1.1	70
26.	The Data Matrix for Problem 1.2	70
27.	One of Two Equally Parsimonious Character Optimizations for Problem 1.2: The Convergence Option	71
28.	One of Two Equally Parsimonious Character Optimizations for Problem 1.2: The Reversal Option	71
29.	The Data Matrix for Problem 1.3	72
30.	Solution 1 of Two Equally Parsimonious Solutions for Problem 1.3	72
31.	Solution 2 of Two Equally Parsimonious Solutions for Problem 1.3	73
32.	The Data Matrix for Problem 1.4	73
33.	Solution 1 of Three Equally Parsimonious Solutions for Problem 1.4	74
34.	Solution 2 of Three Equally Parsimonious Solutions for Problem 1.4	74
35.	Solution 3 of Three Equally Parsimonious Solutions for Problem 1.4	75
36.	The Data Matrix for Problem 1.5	76
37.	Solution 1 of Four Equally Parsimonious Solutions for Problem 1.5	76
38.	Solution 2 of Four Equally Parsimonious Solutions for Problem 1.5	77
39.	Solution 3 of Four Equally Parsimonious Solutions for Problem 1.5	78
40.	Solution 4 of Four Equally Parsimonious Solutions for Problem 1.5	78
41.	The Data Matrix for Problem 2.1	79
42.	The Most Parsimonious Solution for Problem 2.1	79
43.	The Data Matrix for Problem 2.1a	80

List of Figures—Continued

44.	The Most Parsimonious Solution for Problem 2.1a	80
45.	The Data Matrix for Problem 2.2	81
46.	The Most Parsimonious Solution for Problem 2.2	81
47.	The Data Matrix for Problem 2.2a	82
48.	The Most Parsimonious Solution for Problem 2.2a	82
49.	The Data Matrix for Problem 2.3	83
50.	The Most Parsimonious Solution for Problem 2.3	83
51.	The Data Matrix for Problem 2.3a	84
52.	Solution 1 of Two Equally Parsimonious Solutions for Problem 2.3a	84
53.	Solution 2 of Two Equally Parsimonious Solutions for Problem 2.3a	85
54.	The Data Matrix for Problem 2.4	85
55.	Solution 1 of Two Equally Parsimonious Solutions for Problem 2.4	86
56.	Solution 2 of Two Equally Parsimonious Solutions for Problem 2.4	86
57.	The Data Matrix for Problem 2.4a	87
58.	The Most Parsimonious Solution for Problem 2.4a	87
59.	The Data Matrix for Problem 4.3	88
60.	The Most Parsimonious Solution for Problem 4.3	88
61.	The Data Matrix for Problem 5.3	89
62.	The Most Parsimonious Solution for Problem 5.3	89
63.	A Procedural Model of Expert Phylogenetic Tree Construction	108
64.	Model Components Associated With "Look at Single Characters"	112
65.	Model Components Associated With "Compare Characters"	114
66.	Model Components Associated With "Look at Rows"	117
67.	Model Components Associated With "Modify Matrix"	118

List of Figures—Continued

68. Model Components Associated With "Consider Inclusion/Exclusion"	121
69. Model Components Associated With "Consider Parsimony"	126
70. Model Components Associated With "Make Inference About Tree"	128
71. Model Components Associated With "Modify Tree"	131
72. Most Parsimonious Trees as a Subset of Largest Inclusion/Exclusion Trees	165
73. Incomplete Intersection of Largest Inclusion/Exclusion Trees and Most Parsimonious Trees	166
74. Non Intersection of Largest Inclusion/Exclusion and Most Parsimonious Trees	167
75. Reproduction of S7's Paper-and-Pencil Initial Trees and Final Phylogenetic Tree for Problem 5.3	174
76. S8's Solution to Problem 5.3	179
77. Percent of Participants that Found Topologies Plotted Against Signal Index.....	193
78. A Wagner Tree for Problem 1.1 as Constructed by S2.....	205
79. A Model of Desired Performance for Phylogenetic Tree Construction.....	209

CHAPTER I

INTRODUCTION

Teaching of Evolution

Evolution is unquestionably the most important theoretical framework in biology, explaining both the vast diversity and similarities among living and fossil organisms. Dobzhansky (1973) expressed it by saying "Nothing in biology makes sense except in the light of evolution" (p. 125). The change in viewpoint associated with adopting an evolutionary perspective of biology has been compared with the shift from a Ptolemaic to a Copernican system of planetary motion.

Unfortunately evolution is rarely accorded a place in the biology curriculum commensurate with its importance within the discipline (Skoog, 1992). Evolution is often equated with natural selection and is taught from a functional perspective rather than using comparative and historical approaches that are critical for developing an appreciation of the power of evolutionary theory (Mayr, 1988). Evolution is poorly understood and widely disparaged among both teachers and American society at large (Skoog, 1992).

The teaching of evolution represents an important challenge to science educators. Science for All Americans (Rutherford & Ahlgren, 1990), Project 2061 (Clark, 1989) and the report of the Committee on High-School Biology Education (National Research Council, 1990) all highlight the need to place greater emphasis on evolution. Improvements in students' learning of this domain could revolutionize the understanding that they gain from the study of biology.

Problem Solving

The problem-solving tradition in science education is based upon the premise that insight gained from the study of how scientists actually solve scientific problems can be a useful guide to teaching and curriculum development. This tradition in science education is an extension of practice begun in cognitive psychology. Early studies examined content-free problems and led to the development of detailed studies of complex domain-specific problem solving in physics (Chi, Feltovich, & Glaser, 1981) and biology (Collins, 1986). Problem-solving research in biology has focused on transmission genetics, a domain that, unlike most in biology, has been traditionally conceptualized as problem-based (See Stewart and Hafner, 1994). In some advanced courses, especially in population genetics and ecology, problem solving is sometimes used in the context of applying mathematical models to data. In most introductory courses, however, students are expected to learn particular answers rather than explore fundamental questions or develop habits of critical thought (Clark, 1989).

A problem-based approach to the teaching and learning of evolution may offer a number of benefits to students. Stewart (1988) has outlined four classes of potential learning outcomes from the use of problem solving in genetics: (1) the conceptual structure (laws, theories, and their organization) of a particular discipline; (2) problem-solving heuristics that are not specific to a particular discipline; (3) content-specific problem-solving procedures (domain-specific instantiations of general heuristics and problem-solving algorithms specific to the domain); and (4) insight into the nature of science as an intellectual activity (Stewart, 1988). A problem-based approach to the teaching and learning of evolution could be facilitated through the use of a problem-solving environment.

A problem-solving environment is a simulation of a phenomenon of interest, often computer-based, that facilitates students' abilities to gather and analyze data about the phenomenon and to propose and test solutions. Problem-solving environments allow students to address "realistic" problems, that is, problems similar to those faced by practitioners within a domain that require the problem solver to reason from effects to causes. Through the generation of realistic problems, a problem-solving environment can illustrate specific conceptual components and relationships without requiring the student to master complex methodologies. In this way, a computer-based problem-solving environment for phylogenetic analysis offers the potential to improve the teaching and learning of evolutionary biology.

Phylogenetic Investigator

Phylogenetic Investigator (PI) is a program for the Macintosh computer that offers teachers and students a problem-solving environment for addressing problems of phylogenetic analysis using directed search. PI is intended for introductory students in biology at the secondary and undergraduate levels. Students create phylogenetic trees that show the postulated evolutionary pathways that relate fossil and recent species to one another. Transitions between character states are represented on links between species. The transitions explain the pattern of character states that the organisms possess.

PI provides an environment which allows flexibility in creating representations and does not limit the ability of students to think creatively about the nature of phylogeny and the models which underlie it (models of speciation, population genetics, microevolution, hybridization, gradualism, punctuated equilibrium, etc.). PI facilitates teaching using the 3 P's of Stewart and Jungck (1994): problem posing, problem solving, and peer persuasion. This type of problem-based

teaching offers increased potential for students to learn the conceptual structures, domain-specific algorithms and heuristics and to develop insight into the nature of science as an intellectual activity (Stewart, 1988).

Current methods of phylogenetic analysis rely largely on undirected search to constrain possible directed hypotheses (Forey, Humphries, Kitching, Scotland, Siebert, & Williams, 1992). PI, however, was written to engage students in directed search for several reasons. First, directed-search problems are less abstract for beginners than undirected-search problems while still allowing students to explore the basic conceptual framework of phylogenetic analysis. Second, directed-search methods result in phylogenetic histories that are readily interpretable by beginners, while the networks that result from undirected-search are not. Third, directed-search methodologies are more useful for teaching the foundations of phylogenetic inference because they are more clearly related to the Hennigian patterns of argumentation from which cladistic systematics was derived.

Problem Statement

The goal of this study is to begin an examination of problem solving within the domain of evolutionary biology as performed by recognized experts in the field. This study is intended as the first step of a long term investigation to examine the content knowledge, strategic knowledge and organization of knowledge structures required to create phylogenetic trees. This study examines how experts generate trees from coded and polarized data. Subsequent investigations can build upon this by investigating other aspects of phylogenetic analysis and how the products of analysis are compared, brought into agreement, interpreted, and applied.

The purpose of this study is to explain expert performance in the construction of phylogenetic trees. In particular, what are sought are the strategies and heuristics

that experts use. Strategies include the methods, procedures, practices and principles used to solve problems, while heuristics involve more basic reasoning methods that produce plausible results (Darden, 1991). Heuristics often represent particular implementations of strategies that use salient characteristics of the problem to greatly limit the search space required to find a plausible solution, though not necessarily a correct one. Both strategies and heuristics are based on content knowledge, strategic knowledge, and how knowledge is integrated and organized. Therefore, the three research questions to be answered are:

1. What content knowledge do experts possess that enables them to create phylogenetic representations from coded data effectively?
2. What strategic knowledge do experts possess that enables them to create phylogenetic representations from coded data effectively?
3. How is the knowledge of experts organized that enables them to create phylogenetic representations from coded data effectively?

The validity of this study is dependent upon the extent to which Phylogenetic Investigator allows experts to demonstrate the kinds of strategies and heuristics that characterize their expertise. During the development of the research problems and problem-solving environment, biologists who conduct phylogenetic analysis were consulted regarding the appearance and functionality of the program and its presentation of data to insure that it could allow experts to solve problems in a way that would be likely to demonstrate their expertise. In addition to the research questions outlined above, the validity of PI as a tool for phylogenetic tree construction will also be explored with each subject.

Significance

Explaining expert performance in phylogenetic tree construction has theoretical significance within the biological community, the problem-solving research tradition in science education, and the practice of teaching evolution. In addition, there will be practical implications for the improvement of Phylogenetic Investigator, other software that addresses phylogenetic inference, and the development of curricula that address evolution. Finally, this work can serve as the foundation for further research into phylogenetic inference.

Although a number of articles and texts have been written describing methods of phylogenetic tree construction, few have described in detail the methods used to generate the inferences necessary for tree construction. This quote from The Compleat Cladist, a manual of phylogenetic procedures, is characteristic: "A really experienced phylogeneticist ... will 'inspect' the data matrix and produce a first tree based on this general inspection, filtering the data in her mind". (Wiley, Siegal-Causey, Brooks, & Funk, 1991, p. 46) This study can complement and extend previous research describing methods of phylogenetic systematics by investigating the processes by which experienced phylogenetic biologists actually filter data in their minds.

Branching diagrams are used in at least three areas of biology to represent relationships among individuals over time: fate maps for cells during development, genealogical pedigrees, and phylogenetic trees. These areas represent three levels of biological organization in which individuals, whether cells, organisms, or species, have ancestor-descendant relationships resulting in patterns of cascading bifurcations.

Insight into tree construction at one level has the potential of informing the process at other levels of biology as well.

Understanding how experts create phylogenetic representations breaks new ground in problem-solving research. Problem-solving research began with content-free problems, but over time the emphasis has shifted to domain-specific problems. In science, problem-solving research has until recently been applied primarily to areas that have been traditionally conceptualized as problem based, like physics and genetics. This project applies methods from the problem-solving research tradition to an area that has not previously been conceptualized as problem based.

This project has the potential to inform the practice of teaching of evolution. The subject of evolution is controversial among the American public. This controversy is partly due to confusion about the claims of evolutionary theory and partly due to the wholesale rejection of several scientific methodologies, particularly with respect to the dating of fossils (Kitcher, 1982). Improving the way evolution is taught has the potential of enhancing the image that scientific inquiry has among the American public. Even among those who accept evolution, fundamental aspects of phylogeny are widely misunderstood. Gould (1989) described public perception of evolution as being dominated by an iconography of progress showing, in linear form, a marching line of apes to humans. This one image more than any other has entered into the popular psyche as the default appearance of evolution. The teaching of phylogeny has the ability to address this issue by having students construct phylogenies using the same techniques as scientists. Insights gained from this activity could allow students to recognize overly-simplistic and misleading views of evolutionary "progress" and replace linear conceptions of progress with a branching tree structure that more accurately reflects current scientific theory.

Most people are probably aware of the existence of scientific schemes of classification. Most people would recognize the notion that dogs and wolves and coyotes are all members of some class that doesn't include lions and tigers or bears. They probably would also recognize a superclass, mammals, that does include all of these. Few, however, could provide any theoretical basis for why these organisms should be classified together. Evolution provides that basis. Without an understanding of the mechanisms that have produced the diversity of life on earth, the study of classification for most students involves nothing more than memorizing a vocabulary.

During the past thirty years, there has been considerable scientific controversy in systematic biology (Hull, 1988; Scott-Ram, 1990; Ridley, 1986). The history of systematics provides an excellent example for demonstrating to students the processes by which scientific knowledge is constructed and negotiated. In particular, it highlights the relationships between theory and methodology and between theory and data. Phylogenetic systematics can provide a case study for teaching students how the scientific process can be examined and understood.

Mayr (1988) has argued that biology consists of two distinct sub disciplines: functional and evolutionary biology. Functional biology, which deals with questions of proximate causes, has largely been based upon experimental methodologies while evolutionary biology, which deals with questions of ultimate causes, has made extensive use of comparative methodologies. Mayr argues that current teaching practices emphasize functional and experimental biology. Instruction that fails to acknowledge both functional/experimental and comparative/evolutionary components of biology leads to students that construct incomplete and incorrect conclusions about biology. Introducing students to phylogenetic analysis, by which students can make historical inferences based on comparative data, provides opportunities for students to

engage in problems of ultimate causation and may help students construct a more balanced picture of the biological domain.

Problem-solving research has been useful for instructional development in many content areas. A framework for developing instructional systems developed by Reif (1983) uses a rational analysis of problems and studies of expert performance to develop models of desired performance. A study of phylogenetic tree construction and how experts perform this task can contribute to the development of a model of desired performance for phylogenetic analysis which can guide teaching and curriculum development.

This research has practical implications for the continued development of Phylogenetic Investigator. The study of experts can inform the construction of problems that PI should be able to present to students and the improvement of the tools that PI provides for manipulating and representing phylogenetic data. Finally, PI can be modified to facilitate strategies and heuristics used by experts in a form that novice problem solvers can use.

Some of the earliest studies of expert performance in problem solving examined chess grand-masters (Charness, 1989). Since that time, in large part based on insights gained through the study of expertise, computer programs have been developed that play chess as well as all but the very best human players (Charness, 1989). Similarly, for phylogenetic analysis, computer programs are now routinely used to search the problem space for the most parsimonious trees. Insight into how experts create trees may inform the development of improved programs.

This research can inform the implementation of problem-based approaches to teaching evolution. An improved understanding of the nature of phylogenetic problems can help teachers ensure that instruction addresses all of the relevant issues. A model of desired performance can be used to establish goals for student learning

that will result in good problem solving. In addition, this research can provide a knowledge base that teachers can draw from to understand how experts apply components of the model in a range of different circumstances.

What is offered here is a first step that brings a well-grounded tradition—problem-solving research in science education—to a new domain. This work can provide a foundation for subsequent research to build upon. Understanding how experts use and interpret trees would complement and extend the results of this study. A model of novice student performance could be used to complement the model of desired performance described here and potentially improve instructional development. Such a research program has the potential to transform the teaching of evolution.

CHAPTER II

LITERATURE REVIEW

Introduction

Recent efforts to improve the teaching of evolution have largely used a conceptual change model of learning (Posner, Strike, Hewson, & Gertzog, 1982) to address why many students fail to appropriate scientific conceptions of evolution. Several recent efforts are summarized below. They include Bishop and Anderson (1986) who investigated problems in conceptual change related to microevolution, Trowbridge and Wandersee (1994) who applied techniques of concept mapping to improve conceptual change, and Duveen and Solomon (1994) who developed a series of activities to provide a broader cultural context for understanding Darwin's theory of evolution. This study seeks to illuminate a new path to improving the teaching of evolution that uses realistic problem solving to approach macroevolution.

Conceptual change research seeks to explain why many students have difficulty achieving an adequate understanding of a subject when seemingly simple and straightforward explanations have been provided (Posner et al., 1982). The goal of conceptual change research is to elicit students conceptions within a domain and to use this information to create teaching materials and techniques which provide experiences and explanations that directly address the experiences, attitudes and preconceptions which are predominantly responsible for student difficulties. In the teaching of evolution this work has taken the form of describing student conceptions, assessing the effects of instruction, and noting instances of difficulty. Bishop and Anderson (1986) found three areas in which student conceptions clashed with the

scientific conceptions. Students perceived evolutionary change as a single process by which the environment affects individuals. A more scientific conception requires two interlocking processes: one that stochastically increases diversity and the other that non-randomly decreases it. Students also placed little importance on the role of variation within a population—populations were seen as generally homogeneous with change occurring throughout. Finally, students saw evolution as a change of the trait itself, rather than a change in the number of individuals that possess a trait.

Bishop and Anderson (1986) asked students to write scenarios of biological change from an evolutionary standpoint. In order to write these stories correctly, students must use scientific conceptions of evolution, including conceptions of populations, natural selection, and adaptation. Students study each situation provided and invent a story that conforms to a generalized pattern of evolutionary histories. Gould and Lewontin (1978) have criticized this form of explanation for overemphasizing both the importance of natural selection as a means of evolution and the results of evolution as producing structures that are optimally designed for some function. Eldredge and Cracraft (1980) argue that this type of reasoning, when uncritically applied, fails to use historical patterns to evaluate assumptions and notions of process. Rutowski (1992) argues for the need for student reasoning at both microevolutionary and the phylogenetic levels in order for students to develop a balanced picture of pattern and process in evolutionary reasoning.

Concept mapping has also been used to improve students conceptual understanding of evolution. Trowbridge and Wandersee (1994) used concept mapping to identify key places where conceptual change seems not to occur. The goal of concept mapping is to focus student's attention on the relationships between concepts, particularly hierarchical relationships. This approach may help students understand how concepts are related and integrated.

The Science, Technology, and Society studies tradition seeks to elaborate the linkages between the development of technology and scientific theories and the society from which those theories emerge. Duveen and Solomon (1994) have created a role-playing scenario that involves students in the societal matrix of the historical development of the theory of evolution. Students role play a character in a fictitious trial to determine whether On the Origin of Species is blasphemous. The scenario has been designed to illustrate not only the fundamental aspects of evolution and the evidence that supports it, but also key aspects of the scientific tradition: that scientists can hold theories about which they are not entirely sure or that scientists may choose to ignore currently unexplained anomalies rather than rejecting theories out-of-hand. Although it is important to establish the linkages between science and society, phylogenetic inference provides an authentic context where students could be engaged in the type of inquiry that scientists themselves pursue. Realistic problem solving, which models this form of inquiry, offers a greater opportunity for students to understand the value and utility of evolution as a unifying theory in the context of explaining biological phenomena.

A problem-based approach to the teaching of evolution that focuses on the historical patterns of diversity and similarity among living and extinct organisms engages students in scientific reasoning about the nature of evolution in a way that has not been systematically addressed previously. The problem-based approach places students in situations in which they confront realistic problems and solve them using scientific patterns of reasoning. The expression "scientific patterns of reasoning" is used here to mean the methods and habits of thought that make reference to the explanatory models that experienced practitioners within a domain use when solving problems within that domain. In the course of solving problems, students may increase their understanding of: (a) the conceptual nature of the domain; (b) both

general and domain specific heuristics and algorithms; and (c) the nature of science as an intellectual activity (Stewart, 1988). In particular, problems that require students to reason from effects to causes appear to offer the greatest potential for student learning.

Toward a Problem-Based Instructional System

In order to capitalize on the potential learning outcomes that result from realistic problem solving, it is necessary to develop some form of instructional system that is based upon the problem-solving perspective. Reif (1983) has put forward a model for the development of a system of instruction from a problem-solving perspective that requires four components: (1) a model of novice student, (2) a model of desired performance, (3) a model of learning, and (4) a model of teaching. In actuality, these models are not isolated components, but must be overlapping and integrated systems of thought.

A model of desired performance requires an underlying theoretical model that specifies the thought processes and forms of knowledge that lead to desired performance. Studies of expert performance and rational analysis of the task are two components that should be considered in the development of a model of desired performance. Below I have summarized the relevant literature on the study of expertise and elaborate two areas where these studies have been applied in biology: Mendelian genetics and diagrammatic reasoning associated with meiosis. Subsequently, I have provided a brief history of systematics and an overview of phylogenetic inference. Finally, I have outlined features of a theory of expertise for phylogenetic inference that are predicted by the literature.

Expertise

The study of expertise seeks to explain the mechanisms that distinguish outstanding individuals in a field from both other individuals in the field and from the general population. The current tradition in the study of expertise seeks to describe the stable acquired characteristics that enable expert performance by individuals in a domain.

The Study of Expertise

Early scientific studies of expertise sought to explain differences between experts by means of primarily inherited characteristics (Ericsson & Smith, 1991). Methods that have been applied to the investigation of expert performance have included the study of basic cognitive processes, IQ, spatial ability, and personality characteristics. Although some of these differences are strongly correlated with performance among the population at large, few are strongly correlated with performance among experts. This lack of correlation has been used as evidence to support the hypothesis that the characteristics that distinguish outstanding individuals in most subject domains are not solely the result of inheritance.

Current methods of studying expertise, collectively termed the "expertise approach," have been used to describe expertise as the result of acquired skills (Ericsson & Smith, 1991). These skills, typically cognitive, give an expert an improved ability to store and retrieve relevant information and to use that information to draw inferences.

The Expertise Approach

The expertise approach is an attempt to systematically study expert performance. This approach uses three phases: (1) capturing superior performance; (2) analyzing superior performance; and (3) accounting for superior performance (Ericsson & Smith, 1991).

Capturing Superior Performance

Activities in real life are complex and intertwined. The goal of expertise study is to identify recurrent activities related to expert performance that can be reliably reproduced under controlled conditions. In particular, goal-directed activities that yield observable behaviors in response to appropriate stimuli are sought.

A small number of well-defined tasks are used to produce a detailed task analysis for a representative problem in the domain. It is usually not possible to conceptualize these tasks as a representative sample from some population of tasks, as knowledge of complex domains are incomplete and the parameters of a population of tasks are, therefore, unknown.

Little attention has been given to the critical process of selecting appropriate stimuli to elicit expert performance. There is evidence that stimuli which are overly abstracted from an expert's range of experience will lead to the application of sub-optimal strategies that will be subject to severe practice effects (Gillhooly, Wood, Kinnear, & Green, 1988; Ericsson & Smith, 1991). As long as practice effects are comparatively small and the expert performance remains superior to that of novices even after extended practice, Ericsson and Smith (1991) argue that superior expert performance can still be adequately captured.

Concurrent (think-aloud) verbalizations collected during expert problem solving coupled with observations of moves through a search space (task environment) have been successful in the past in showing what information is being processed and which types of information processing activities are being employed (Ericsson & Simon, 1980, 1993). There is strong evidence that this form of data gathering does not have significant effects on the nature of the activity being performed, although it may result in slower problem solving when compared with silent problem solving.

Observations that can be drawn from the verbalizations include: the physical objects and mental entities being considered by the participant, relationships between and among those objects and entities, and the causal inferences made during the course of problem solving. The relationship between physical objects and mental entities has been described as two related spaces (Kotovsky & Fallside, 1989). Physical objects exist in the search space or task environment: the external problem environment and the manipulations that can occur within it. Mental entities are part of a problem space: the states of knowledge and the operations that advance the current state of information into a new state. It is necessary to perform analyses of several protocols on similar tasks in order to identify the consistent patterns of cognitive processes and retrieved knowledge. Further analysis can indicate the kinds of strategies people use: was the person working forward from the problem statement, backward from the goal, or were they using some other strategy?

Analyzing Expert Performance

A task analysis, for relating the verbalized information to the underlying cognitive processes should be performed prior to data collection. This analysis can provide confidence that the task, as constructed, has the potential to answer the

research questions. Larkin and Rainard (1984) suggest constructing a model composed of condition-action statements that could explain performance of the task. In addition, the task analysis can specify a number of different sequences of processing steps that could yield the correct answer. In some domains this can be exhaustively described, although in more complex domains only a subset of possible pathways can be considered. Using the subjects' verbalizations, one can determine only that the verbalized information was accessed. During data analysis, the task analysis relates the verbalizations to underlying cognitive processes (Ericsson & Simon, 1993). Verbalizations which do not match the task analysis can provide insight into processes not yet described.

Accounting for Superior Performance

Superior performance in a given domain reflects processes and knowledge specific to that domain. These processes imply the presence of cognitive structures that are specific to the relevant task domains. Chase and Simon (1973) proposed that the main difference between novices and experts was related to their immediate access to relevant knowledge. Recognized configurations in the search space are stored and retrieved in "chunks" that organize a large amount of information into a compact form which requires little mental processing and can be of immediate use.

Recent studies of expert performance have focused on practice and learning, memory functioning, and experts' ability to plan and reason (Ericsson & Smith, 1991). From this body of research has emerged a pattern of consistent features of expertise.

The Nature of Expertise

Experts access relevant knowledge efficiently which facilitates the recognition of useful regularities and inconsistencies. On presentation, problem information is integrated with the relevant domain knowledge. Expert performers tend to retrieve a solution method as part of the immediate comprehension of the task.

Anzai (1991) noted that solving physics problems requires: (a) recognizing the underlying structure of the problems, (b) generating representations appropriate for discovering solutions, and (c) making inferences regarding the generated representations. Recognition involves theoretical knowledge of physics, generation is often accompanied by abstract diagrams, and inferences are made by efficient strategies such as working forward. These components of expertise are not unique to physics (Anzai, 1991).

Problem Recognition and Data Redescription

Experts describe problems in terms of content principles rather than surface features (Chi, Feltovich, & Glaser, 1981). Organization of experts' knowledge affects how they redescribe problems. Experts create physics problem descriptions at several levels: English (as presented in the text), basic (facts and relationships that the solver views as important), science (conversion of facts and relationships to scientific concepts and principles), and computational (formulas required to operate on these concepts and principles) (Larkin & Reif, 1979; Larkin & Rainard, 1984). One important component of problem representation is the generation of a problem diagram. In physics, this might take the form of a drawing which illustrates all of the forces acting on a body to predict its future direction and speed.

Larkin (1983) described schemata which problem-solvers use to produce representations of physics problems. These schemata contain two types of rules: construction rules and extension rules. Construction rules allow the translation between the original problem representation (in everyday terms) and the terminology of physical entities. Extension rules allow the creation of new entities within a physical representation based on relationships between entities within the representation. Continuing the example from above, construction rules govern translating the forces acting on the body into vectors and extension rules allow the generation of the sum vector to predict future direction and speed.

The representations created according to these schemata have characteristics that distinguish them from naive representations: the entities are technical with meaning only within the scientific domain; the inferencing rules are qualitative, time independent and redundant; the representation is closely associated with fundamental principles of the domain; and the properties of entities are localized.

When drawn by experts, diagrams tend to be principle oriented abstractions of physical objects and their relations that are substantially different from realistic drawings. Both self-generated representations and representations generated by others can be equally useful for making inferences (Anzai, 1991). In both the recognition and generation processes, representations may include appropriate information for solving problems (Larkin, 1991).

Diagrams have been suggested to function by encoding problem schemata and retrieval cues for recognition (Gick & Holyoak, 1983). In addition, they may provide good spatial cues for efficient inference making either by grouping together all the information needed to make inferences or by organizing information by location facilitating its retrieval (Larkin & Simon, 1987).

Larkin has extended this work by defining a realm of problem solving she terms display-based (Larkin, 1989). In this form of problem solving, the current state of the problem solution represents its main display of the problem. This form of problem solving reduces the need for short term memory by storing products of subcomponents of the problem-solving process directly within the developing solution.

Coordination of Recognition and Inference

Much literature on expert problem solving suggests that recognition and inference are the two principal factors that dominate expert performance when the activity can be considered as seeking goals under constraints (Greeno & Simon, 1988; Anzai, 1991). In these contexts, experts are more likely to use forward-working knowledge producing strategies rather than end-means (Larkin & Rainard, 1984). In spite of the generality of the inference component in different domains, the methods of inference are domain specific. Formation of inference strategies in a domain is affected by the structures of the tasks and the knowledge of the problem-solvers. Although forward-working strategies are useful in physics, where problems tend to be strongly constrained, in other domains working backwards from goals may be more useful (Anzai, 1991).

A Study of Experts in Genetics

Collins (1986) combined studies of experts solving transmission genetics problems and a rational analysis of these problems to develop models of desired performance. Experts used Genetics Construction Kit (Jungck & Calley, 1984) to pose and solve several transmission genetics problems. Think-aloud protocols were analyzed to identify the strategic knowledge that experts were using during three

phases of problem solving: data redescription, solution synthesis, and solution assessment.

In transmission genetics, data redescription involves specifying the number of traits and their names, the number of variations per trait and their names, and unusual distributions of variations with respect to traits. There are four occasions during the problem-solving process in transmission genetics that involve data redescription: (1) at the beginning to redescribe initial data and form a working hypothesis, (2) during problem solving to consider alternate hypotheses, (3) after confirmation of inheritance patterns, and (4) at the beginning of the second part of a two-stage problem. Solution synthesis in transmission genetics is synonymous with hypothesis testing. Two types of hypotheses are involved in solving these problems: general hypotheses about inheritance patterns and modifiers and specific hypotheses about the results of data producing crosses. Solution assessment strategies in transmission genetics involve confirmation of hypotheses through additional crosses or mathematical tests.

Collins (1986) contrasted genetics and physics problems and problem solving (see Table 1). She found that experts' content knowledge of transmission genetics is organized around the deeper structure of the domain and does not rely on surface features of problems. She also found that experts have ancillary knowledge that determines when to apply conceptual knowledge to a problem. Collins concluded that the strategic knowledge of genetics experts, like physics experts, is organized into the three categories of data redescription, solution synthesis, and solution assessment. Expert strategies in solving transmission genetics problems included using an initial problem redescription to identify essential features of the problem and using knowledge-producing strategies to increase their understanding of the problems. In spite of differences between specifics of physics and genetics problems and problem solving, patterns of expertise are consistent across these two domains.

Table 1
A Comparison of Characteristics of Physics and Genetics Problems and Problem-Solving

Characteristic	Physics	Genetics
Problems		
Data	<ul style="list-style-type: none"> • limited to what is given in problem 	<ul style="list-style-type: none"> • continuous production is possible
Solution	<ul style="list-style-type: none"> • requires a mathematical formula • has a numerical value 	<ul style="list-style-type: none"> • no mathematical formula applies • is not numerical
Nature	<ul style="list-style-type: none"> • qualitative only in problem statement 	<ul style="list-style-type: none"> • qualitative descriptions throughout problem
Data Redescription		
Number	<ul style="list-style-type: none"> • once 	<ul style="list-style-type: none"> • many times
When	<ul style="list-style-type: none"> • beginning of problem 	<ul style="list-style-type: none"> • with each hypothesis
What	<ul style="list-style-type: none"> • from qualitative statement in text through four levels to computational statement 	<ul style="list-style-type: none"> • qualitative statement of names of traits & variations, and quantitative statement of numbers of traits, variations, and classes

Table 1—Continued

Characteristic	Physics	Genetics
Solution Synthesis		
Types	<ul style="list-style-type: none"> • means/end analysis • knowledge producing • setting subgoals 	<ul style="list-style-type: none"> • hypothesis testing which is knowledge producing • setting subgoals
Requires	<ul style="list-style-type: none"> • computational formula 	<ul style="list-style-type: none"> • definitive cross
Solution Assessment		
Consists of	<ul style="list-style-type: none"> • numerical value consistent with problem statement with • interpretability and completeness • internal and external consistency • optimality 	<ul style="list-style-type: none"> • statement of the inheritance pattern • genotype-to-phenotype match • confirmation • test of modifiers

Note. From Strategic Knowledge Required for Desired Performance in Solving Transmission Genetics Problems. (p. 229 and 236), by A. Collins, 1986, Doctoral Dissertation, University of Wisconsin, Madison. Copyright 1986 by Angelo Collins. Adapted with permission.

Diagrammatic Reasoning in Meiosis

Kindfield (1994a) compared expert and novice abilities to create and use representations of meiosis. She has argued that a fundamental interconnection exists between meiosis knowledge and representational skill. Experts routinely generate and use finely-tuned diagrams in which relevant knowledge can be immediately applied to the task at hand. Domain knowledge and pictorial ability coevolve: an increase in knowledge of meiosis increases representational ability and the development of pictorial skill aids in the construction of accurate mental models. Diagrams appear to serve as information storage devices that free working memory for other cognitive tasks related to problem solving.

Similar results had previously been obtained in research into physics problem solving (Larkin & Simon, 1987; Larkin, 1989). Diagrams serve a heuristic function by grouping information by location; aid making inferences regarding features and relationships essential to problem solving; and help the solver remember related information not initially represented. Kindfield's research extends previous studies by describing how diagrams may also play a role in building knowledge about meiosis (Kindfield, 1994b).

A Brief History of Systematics

Phylogenetic analysis has undergone a series of upheavals over the last 30 years (Hull, 1988). Prior to the 1960's, evolutionary taxonomy was the dominant school of thought in systematics with methodologies that were rooted in the neoDarwinian tradition. Evolutionary taxonomy proved unable to respond to criticisms that systematics and resulting classifications were based upon circular logic (Brooks & McLennan, 1991).

The phenetics movement began with taxonomists who were frustrated at the lack of a solid theoretical base upon which taxonomy could be grounded. Numerical taxonomists solved problems of circularity by divorcing systematics from evolutionary theory and classified organisms based solely on similarity. Scott-Ram (1990) identifies J. S. L. Gilmour as the first taxonomist to outline what would become the framework for numerical taxonomy. Sokal and Sneath (1963) elaborated in greatest detail what has been called the phenetic or numerical taxonomy tradition. Pheneticists believed that an objective system of classification could be constructed based purely upon morphological difference by using large numbers of characters. The phenetics movement was also aided by pioneering use of computers to manage the data necessary to consider large numbers of characters (Scott-Ram, 1990).

A German work originally published in 1950 became available in English translation in 1966 and was subsequently adopted by a group of taxonomists who were dissatisfied both with evolutionary taxonomy, for its circularity, and with numerical taxonomy, for its lack of any evolutionary theoretical framework. Hennig (1966) proposed a non-circular methodology for establishing phylogenetic relationships. This practice, called cladistics in English, soon developed a wide following. Cladistics is concerned only with monophyletic groups. These groups are defined as all of the descendant taxa of single ancestral species. Cladists criticized both phenetics and evolutionary taxonomy for constructing groups based on overall similarity rather than genealogy. Taxonomy based on overall similarity will sometimes recognize paraphyletic and polyphyletic groups. Paraphyletic groups are an incomplete set of descendants of a common ancestor while polyphyletic groups contain additional species. (See Appendix A for a summary of important concepts from phylogenetic systematics).

Classifications based solely on genealogy gained wide acceptance during the 1980's (Scott-Ram, 1990). During this period there was renewed interest on the relationship between classifications and evolutionary theory. Differences over this question ultimately led to a split among the cladists. A small number led by Nelson, Platnick, and Patterson formed what has been called transformed cladistics, a tradition that uses cladistic methodologies, but which seeks to eliminate influences on classification which derive from interpretation based on evolutionary theory (Hull, 1988). Critics of transformed cladistics (Ridley, 1986; Scott-Ram, 1990) argue both that classifications cannot be divorced from all theoretical influences and that to do so would be undesirable.

This study draws its primary influences from the "evolutionary" cladistic school of phylogenetic systematics and all of the experts studied are from this school. A potentially interesting topic for further study would be to investigate whether and how experts from different schools construct, use, and interpret phylogenetic trees differently.

Phylogenetic Inference

This section provides an introduction to phylogenetic inference that would be appropriate for a biologist or teacher of biology: someone who is already familiar with the concepts and terms of evolutionary biology and systematics. Appendix A contains a summary of the concepts and terminology that are used in this section as an introduction and reference for other readers.

Phylogenetic inference is conducted in order to answer questions about relationships among taxa and characters. Constructing a phylogenetic tree creates a representation of the data that greatly increases humans' abilities to make inferences about the data that would otherwise be difficult to derive. Table 2 presents a series of

Table 2
Methods Necessary for Answering Phylogenetic Questions

Question	Form of Data	
	Phylogenetic Tree	Data Matrix
Is a character a uniquely-apomorphic or a shared-apomorphic character?	Find unique characters between a species that possesses it and its ancestor. Shared-apomorphic characters are found below common ancestors	Perform phylogenetic inference
Has a given species been previously described or is it unknown?	Look at the branches sequentially and consider whether the species has each apomorphy until reaching a taxon or an inconsistency	Compare a row of the matrix with all other rows
Which groups of species represent evolutionarily complete units (clades)?	Identify all taxa descended from postulated ancestor	Perform phylogenetic inference
How closely related are two or more species?	Find the most recent common ancestor and count the number of steps (character transitions) between taxa	Perform phylogenetic inference
Given a species, what other species may have the same or similar characteristics?	Look at sister taxa to estimate shared properties	Perform phylogenetic inference
When did a character change in state (in relative terms)?	Find the transition on the tree	Perform phylogenetic inference
What was the order of evolutionary change?	Follow the links between taxa and ancestors	Perform phylogenetic inference

Table 2—Continued

Question	Form of Data	
	Phylogenetic Tree	Data Matrix
What similarities and difference exist among patterns of evolutionary change?	Superimpose or compare trees	Perform phylogenetic inference

biological questions asked by systematists and the inferencing methods necessary to answer the question based on using a phylogenetic tree or the data matrix of characters from which a tree would be constructed.

In most of the examples, a process of phylogenetic inference is required to draw the conclusions. This means that in order to arrive at a solution using the data matrix, it is necessary to distinguish among homologous and non-homologous characters, define orders of character transitions, and consider distribution of homoplasious characters. In other words, it is necessary to go through all of the steps involved in the construction of a phylogenetic tree. Individual characters and taxa can only be assessed in relationship to others in the context of a hypothesis or hypotheses that takes all of the characters into account. Phylogenetic trees are a diagrammatic form of representing and preserving these hypotheses.

Principles of Phylogenetic Inference

Hennig (1966) made the observation that shared characteristics resulting from unique common ancestry (homologies) should covary within a group whereas characteristics shared due to evolutionary reversal or convergence (homoplasies) should not. He proposed that by making the initial assumption that all similarities

were homologies, characters that failed to covary during construction of a phylogenetic tree (homoplasious characters) could be detected and eliminated.

Phylogenetic analysis seeks to recognize the genealogical relationships that exist among and between living organisms and/or fossils. These relationships are inferred from character distribution data by means of assumptions of evolutionary process (Sober, 1988). Species and groups of species are described as sister taxa descended from a unique common ancestor based on shared apomorphic characteristics. These statements of common ancestry, or sister-group relationship, are represented in the form of a phylogenetic tree. Phylogenetic trees are generated based on the most "parsimonious" arrangement of character transitions between plesiomorphic and apomorphic characters using any number of methods of determining parsimony.

The problem-solving process can be delimited into approximately four stages. First a group of species of interest that is believed to be monophyletic (the "ingroup") and, usually, at least two closely related taxa (the "outgroup") must be identified . Second, these groups are studied to find characters which show significant interspecific variation. Third, a determination must be made of which character states will be considered plesiomorphic. Fourth, shared apomorphic states among the species of the ingroup are used to construct an evolutionary pathway that requires the least number of transitions (based on a variety of measures of parsimony).

Hennig described phylogenetic inference as a directed search. The search begins with a clear identification of plesiomorphic and apomorphic conditions and a tree is constructed based on those assumptions. More recently, phylogenetic methodologies have begun with an undirected search where a network is constructed that organizes the taxa based on unpolarized characters. A root can then be added anywhere in the network that polarizes all of the characters. Although this has proven

to be a useful exploratory tool for phylogenetic inference, I have chosen not to incorporate this innovation for several reasons. First, it is recent and not universally accepted. Second, it adds a layer of complexity which is not essential for addressing the fundamental conceptions of phylogenetic inference. Third, it results in hypotheses which are not historical, which is potentially confusing to novice students.

Selection of Ingroup and Outgroup

Identification of the ingroup is usually determined by a systematist who begins with a particular group of problem taxa in mind. Usually, it is assumed that the larger taxa (e.g. genera) are already organized with respect to monophyly (Eldredge & Cracraft, 1980) and that the goal of analysis will be to establish the relationships within that monophyletic group. If these relationships are uncertain, a lower-level study may be undertaken first to resolve uncertainty about the ingroup. Lower-level studies often use large numbers of taxa to look for groups that appear to be monophyletic (Stevens, 1991).

Once the ingroup has been defined, the search for an appropriate outgroup is conducted. The outgroup is used to provide evidence of the states of characters prior to the common ancestor of the ingroup. The most desirable outgroup consists of the taxa most closely related to the ingroup. In the event that these relationships are unknown, any closely related group of species not within the ingroup can be used.

Selection of Characters and States

Initially, characters are surveyed within the ingroup and the outgroup with the only requirement being that the character can be recognized when it occurs (Stevens, 1991). Discrete characters are handled most easily by phylogenetic analysis. Continuous quantitative characters can be also used, although no consensus yet exists

on how to code them (Stevens, 1991). Methods for delimiting continuous characters in discrete codes vary depending on the research tradition of the systematist and the problem: some have argued for exclusion of these characters, others for arbitrary delimitation, others for using mathematical rules for delimitation, and yet others for the use of continuous coding scales (Stevens, 1991).

Determination of Polarity

Once states have been defined for characters, their polarity must be considered: which state is to be treated as plesiomorphic and which, apomorphic? The most important methods for determining the polarity of character states are outgroup, paleontological, and ontogenetic comparison. Each method has its strengths and weaknesses. Each can explain certain types of data and each has methods for explaining conflicting data. For all of the methods, conflicting data can be explained through the use of homoplasy (seemingly equivalent characters can be defined as non-equivalent).

Outgroup Method

The outgroup method assumes that homologous transformations result in variation in characters found only within the ingroup. De Queiroz (1985) writes:

Given the existence of a monophyletic group within which occur alternative characters, a phylogenetic character transformation must have occurred within the group. (This logic only applies to cases in which the characters vary among the taxa whose relationships are being investigated, not within them) Thus the character found both inside and outside of the group is plesiomorphic; the character found only inside of the group is derived. (pp. 285-286).

The outgroup method can account for conflicting data by reevaluating whether some outgroups should be considered ingroups or vice versa. The key to successful use of the outgroup method is to have well resolved groups.

Paleontological Method

The paleontological methods uses data from fossil to determine the polarity of characters. "Given alternative homologous characters whose existence is documented in fossils, the character found in the oldest fossil is plesiomorphic, the alternative derived [i.e., apomorphic]." (de Queiroz, 1985, p. 286).

The paleontological method can account for conflicting data through appeals to the incompleteness of the fossil record. The oldest fossil in the fossil record may not represent a direct ancestor of the ingroup. Fossils can greatly improve the resolution of plesiomorphic character states if the fossils are close in temporal position to ancestors of recent species and if a significant percentage of the characters can be unambiguously coded (Huelsenbeck, 1991).

Ontogenetic Method

The origin of ontogenetic comparison is found in von Baer's second law of development which states that less general characters are derived from the most general ones (Gould, 1977; de Queiroz, 1985). Nelson (1978) elaborated this principle calling it the biogenetic law: "Given an ontogenetic character transformation, from a character observed to be more general to a character observed to be less general, the more general character is primitive [ancestral], the less general advanced [derived]" (p. 327). De Queiroz (1985) has extended this principle to a wider range of cases: "given that ancestral characters are retained in descendant ontogenies, ancestral characters are more general than derived characters" (p. 289). Related to this methodology, there has been extensive discussion regarding the use of ontogenetic transformations as characters, rather than for polarity determination and

questioning the appropriateness of these factors to serve both as characters and as a method for determining polarity (de Queiroz, 1985).

Phylogenetic Tree Construction

The goal of phylogenetic analysis is to create a branching diagram, called a phylogenetic tree, that requires the fewest number of changes in state, i.e. is most parsimonious, in order to explain the distribution of apomorphic character states among the taxa under consideration. Phylogenetic tree construction results from the interaction between the methods used to generate possible trees and the criteria used to determine which trees are most parsimonious.

The goal of phylogenetic tree construction is to postulate or recognize common ancestors of recent and fossil species. This process is based on the assumption that there is only one true phylogeny (Brooks & McLennan, 1991). Genealogical descent produces a hierarchy among recent and fossil organisms at least parts of which are discoverable and effectively representable by a branching diagram (Forey et al., 1992).

The primary mechanism of the tree construction process is Hennig's auxiliary principle: Never assume parallel or convergent evolution; always assume homology in the absence of contrary evidence (Wiley et al., 1991). Each character potentially defines two homologous groups: general homologies which share a character in the plesiomorphic state and special homologies which share it in the apomorphic state. Only special homologies, groups defined by shared apomorphic characters, are used as evidence of common ancestry or relationship.

All homologies should covary with each other and with evolution (Forey et al., 1992; Brooks & McLennan, 1991). Character congruence is the decisive criterion for distinguishing homology (synapomorphy) from non-homology (homoplasy)

(Brooks & McLennan, 1991). Homoplasies—characters shared in the apomorphic state for reasons other than homology—should be random and should, therefore, covary neither with each other nor with evolution.

The groupings defined by two characters can be combined into a single hypothesis of relationship if that information allows for the complete inclusion or the complete exclusion of groups that were formed by the characters (Brooks & McLennan, 1991). Incomplete overlap of groupings leads to the generation of two or more hypotheses of relationship, since the information cannot be directly combined into a single hypothesis.

Parsimony is the principle invoked by systematists to argue that given two competing phylogenetic hypotheses, the simpler should be chosen. Parsimony is the assumption of homology whenever possible and application of homoplasy only when necessary. It is only through the assumption of homology that similarities among organisms can be explained (Farris, 1982). Without this assumption, it is impossible to resolve higher taxa. The goal of the use of parsimony is to maximize character congruence (Brooks & McLennan, 1991).

A variety of forms of parsimony exist which correspond to various algorithms for determining parsimony. Wagner parsimony (Kluge & Farris, 1969) uses ordered characters (change from one character to another implies change through intervening characters in the series) which can reverse. Fitch parsimony (Fitch, 1971) uses unordered characters that can be allowed to reverse. Dollo parsimony (Farris, 1977) does not permit parallelisms and allows only single reversals. Each of these forms of parsimony makes particular assumptions about the nature of evolution.

Toward a Theory of Expertise in Phylogenetic Analysis

If experience from the study of expertise in other domains is transferable to the domain of phylogenetic analysis, then some aspects of expertise in this domain can be predicted. Phylogenetic analysis represents an expert methodology for creating representations that are appropriate for making inferences about the data and constructing solutions. Experts should have content knowledge that can readily be brought to bear to recognize patterns and note inconsistencies in the data. Experts should have strategies and heuristics that will allow recognizing the underlying structure of the problems (distinguishing homologous and analogous structures, assigning character polarity, recognizing synapomorphies and homoplasies), generating representations appropriate for discovering solutions (creating parsimonious trees), and making inferences regarding the generated representations (recognizing key differences between competing hypotheses). Finally, experts should also be able to organize their knowledge in ways that allow for efficient information processing.

CHAPTER III

METHODOLOGY

Introduction

The purpose of this research was to describe the knowledge, both content and strategic, and its organization, which experts use to construct phylogenetic trees from coded and polarized data. To achieve this goal, an analysis of the task of phylogenetic tree construction was performed. This analysis was used as a foundation to develop a computer-based problem-solving environment, a rational analysis of phylogenetic tree construction, and two series of problems: a set of model phylogenetic problems and a set of research problems.

Subsequently, a group of expert participants were recruited to participate in the study by thinking aloud while solving the research problems. The think-aloud protocols and the recorded actions from the problem-solving environment were collected along with all notes and drawings. The rational analysis and the expert studies were synthesized to develop a descriptive procedural model of expert performance for phylogenetic tree construction.

Analysis of Task

The initial task analysis was performed to guide the development of a problem-solving environment that could be used to study phylogenetic tree construction for teaching and research. The first step was to decompose phylogenetic diagrams into their constituent parts. Eldredge and Cracraft (1980) was used as a guide to define the objects and actions needed to create cladograms and phylogenetic

trees. A broad reading of the literature led to the conclusion that binary characters were often used by systematists to illustrate conceptual problems and techniques. Finally, during the development of the problem-solving environment, systematic biologists were consulted to evaluate the reasonableness of the form of the problems and the nature of the resultant solutions. None of the biologists who were consulted about the design of the environment and the problems were participants in the research.

The Problem-Solving Environment

Scott-Ram (1990) claimed that in order for phylogenetic inference to be meaningful:

a model of the process of evolution must be constructed, one which is expressed in the form of a phylogenetic tree (which incorporates a time scale). With the subsequent conversion of the tree into a classification, various rules and criteria are adopted which retain the fundamentals of the model of process. (p. 12)

Below is outlined a process model of evolution that is consistent with diagrams created by Phylogenetic Investigator. The model presented here, though general, provides evidence that the environment created is sufficient to represent phylogenetic hypotheses in general.

A Process Model of Evolution

Evolutionary theory is composed of families of interacting models (Thompson, 1989). The two most important families can be termed microevolution, which is concerned with within-species phenomena and macroevolution, which is concerned with among-species phenomena. Microevolution reflects the population genetics of a single species while random mutation increases diversity of all characters and natural selection decreases diversity in a non-random fashion.

Microevolution can be represented in PI with two nodes, at different periods in time and at different points along the scale of morphological divergence, that are connected by a link. Transitions, representing characters which have changed, can be placed along the link that connects the nodes.

Macroevolution occurs when one species gives rise to two (or more) species. Speciation requires a process of population splitting that involves geographic and demographic factors in conjunction with a novel adaptive response (Carson 1985). Carson (1985) argues that founder events and hybridization disequilibrate a polygenic balance that exists within species (Carson, 1975; Carson, 1985). During the process of re-equilibration, two or more species may result through a selective process of reorganization that acts to constrain diversity with respect to sexual reproduction or non-sexual aspects of the environment. Other models have been suggested to explain speciation events, including disruptive selection, parapatric speciation, allopatric speciation, etc. (Eldredge & Cracraft, 1980; Brooks & McLennan, 1991). Macroevolution can be represented by linking two (or more) nodes to a node located earlier along the temporal scale. The arrangement of the later nodes along the axis of morphological divergence determines whether both nodes are perceived to have diverged equally from the originating node or whether one node has retained the plesiomorphic characteristics to a greater extent than the other. Transitions, representing characters which have changed, can be placed along the links that connect the nodes.

Although micro- and macro-evolution are hierarchically related, one is not reducible to the other (Eldredge & Cracraft, 1980; Brooks & McLennan, 1991) (See Brooks & Wiley, 1986, 1988 for an attempt). Macroevolution defines the limits within which microevolution takes place and microevolutionary processes modify the constituents that are acted upon by macroevolutionary processes.

Phylogenetic Investigator

Although there are a number of computer programs for analyzing and reporting phylogenetic or cladistic data, I elected to create a new program, Phylogenetic Investigator (PI), rather than using one of the pre-existing programs, for several reasons. PI is designed to facilitate creative problem solving in phylogenetic analysis for the purpose of teaching and learning phylogenetic inference. Users can identify characters and states, polarize characters, and engage in directed-search phylogenetic tree construction. PI also allows users to (a) make inferences and represent them one step at a time, (b) vary representational features of their trees (such as angle of divergence and time between speciation events), (c) create reticulate tree patterns, and (d) view all of the character transformations at one time. In addition, PI can generate plausible data stochastically for modeling and practicing tree construction. Finally, PI can record the actions made in the tree construction process. No pre-existing program had any of these features.

PI produces diagrammatic hypotheses that illustrate the pattern of microevolutionary and macroevolutionary events. Hypotheses as constructed by PI are consistent with the model of evolution described above. The goal has been to create a flexible environment in which students can explore the full range of phylogenetic representation. Table 3 describes each component of the model of evolution in the form in which it is used in PI.

The Design and Features of Phylogenetic Investigator

The intent and design of Phylogenetic Investigator distinguishes it from other programs used for phylogenetic analysis. MacClade 2.1 (Maddison & Maddison,

Table 3

Objects, States, and Processes That Underlie Phylogenetic Investigator

Objects	
Monophyletic Taxon:	A species and all of its known descendant species. (Represented in PI as an ancestral taxon linked with all descendant taxa.)
Species:	A collection of organisms that maintain a common equilibrated genetic system. (Represented in PI by a Node and the Link prior to that node.)
Character:	A feature of a species that varies in form among other species. (Represented in PI by a Character.)
Relationship:	A line of continuity that joins an ancestral species with a descendant species. (Represented in PI by a Link.)
States	
Recent:	A species described based on collected specimens. (Represented in PI by the letter "R" in the taxon designation).
Fossil:	A species that is described based on fossil evidence. (Represented by an PI by the letter "F" in the taxon designation).
Postulated:	A hypothetical species to explain common ancestry. (Represented by an PI by the letter "P" in the taxon designation).
Plesiomorphic:	The preexisting, ancestral condition of a character. (Represented in PI by the code "0").
Apomorphic:	The evolutionarily-novel condition of a character. (Represented in PI by the code "1").

Table 3—Continued

Processes	
Time:	Although taxa are identified with particular points in time (recent or at some point in the past (fossil)), they are assumed to have maintained an identity through the passage of time that can be represented by relationships and postulated taxa for which direct evidence does not exist. (Represented in PI by the vertical axis).
Extinction:	The process by which species die out: A species always has an ancestor, but may have no descendants. (Represented in PI by terminal fossil taxa).
Microevolution:	The characteristics of species change over time. (Represented in PI by the horizontal axis of morphological change, transitions on links, and the slope of links).
Speciation:	Species can give rise to two (or more) descendants. (Represented in PI by branching points).

1989) is used as an example in the following sections to illustrate how a research-oriented program design may not fulfill the needs of beginning learners. Newer versions of MacClade and a variety of other programs for phylogenetic analysis that are similar in intent to MacClade (e.g. PAUP (Swofford, 1991), PHYLIP (Felsenstein, 1993), and others) have similar problems when used with beginners. Phylogenetic Investigator and MacClade were also conceived for different purposes. The design of each program reflects those purposes and results in differences between a user's problem-solving processes and products.

Processes of Phylogenetic Problem Solving

Researchers often work with a family of related problems for which they have previously considered the relevant methodological issues and assumptions.

Researchers, therefore, can solve subsequent problems within that family without revisiting those concerns for every problem. Research programs for phylogenetic analysis are constructed to take advantage of this fact, which saves time and makes them more useful. Two ways in which research programs embody assumptions are the starting point for problem solving and the available tools for manipulating and analyzing data.

Starting Point

MacClade allows a researcher to begin with a default hypothesis in the general format which the finished hypothesis will take. The default hypothesis generated by MacClade 2.1 is a random arrangement of the taxa connected by an arbitrary system of postulated ancestral relationships. MacClade will generate such a hypothesis even in the absence of any characters. For an experienced problem-solver who is already aware of the extent of the search space (the range of acceptable solutions), this function of MacClade is a useful time-saving feature, but for beginning learners it bypasses important aspects of the problem-solving process which students should confront. During the course of problem solving students should ask themselves at least three fundamental questions: What are good hypotheses like?; Where do hypotheses come from?; and What is the evidence that supports a hypothesis?

What Are Good Hypotheses Like? At the beginning of the problem-solving process, students need to question what form a completed hypothesis should take. Much of the difficulty in problem solving involves determining what the goal is and what form the solution will take. This is a fundamental question that should be addressed by examining good hypotheses or having an instructor demonstrate the

process of creating one. Beginning with a default hypothesis allows students to avoid asking this question since the basic form has already been established.

Where Do Hypotheses Come From? When faced with the problem of creating a hypothesis, students must consider what sources of information are necessary and useful to guide action. There are a number of factors that can be considered in seeking a beginning point for the development of a hypothesis: What is the range of plausible hypotheses? What are the constraints that limit the number of possible hypotheses? Are there patterns or groups of hypotheses? When students are provided with a default hypothesis, these issues can be largely ignored.

What Evidence Supports a Hypothesis? While constructing hypotheses, students must ask at each step of the process whether there is sufficient evidence to warrant adding another link, node, or transition. If a default hypothesis is given to them in the absence of any evidence, there is little concern for this fundamental issue.

PI does not provide a default hypothesis. This challenges students at every stage of the problem-solving process to reflect on these three questions and develop self-monitoring skills to guide the creation of phylogenetic hypotheses. Students must decide for themselves what general and specific form the hypothesis will take and what evidence is necessary to justify each aspect of the development of their hypothesis.

Available Tools

In addition, although less true of MacClade 2.1, other programs often have features that are superfluous for the types of problems most introductory students face. Many of the tools, that may be useful to a practitioner, conceal fundamental conceptual aspects of the problem-solving domain from the user. Algorithms that

search the problem space for the most parsimonious trees are often based on controversial assumptions (see quote by Felsenstein below) and use complex mathematics that introductory students are unlikely to know in detail. These algorithms can yield unexpected and contradictory results in some circumstances (Platnick, Griswold, & Coddington, 1991; Lorenzen & Sieg, 1991; Haszprunar, 1992; Meier & Whiting, 1992; Lorenzen, 1992), a fact that introductory students may not know. Felsenstein (1993) expressed it this way:

In most cases my assertions about what are the assumptions of these methods are challenged by others, whose papers I also cite at that point. Personally, I believe that they are wrong and I am right. I must emphasize the importance of understanding the assumptions underlying the methods you are using. No matter how fancy the algorithms, how maximum the likelihood or how minimum the number of steps, your results can only be as good as the correspondence between biological reality and your assumptions! (discrete.doc p. 3)

Even fairly innocuous tools, like the Ladderize command of MacClade (see below), can act to constrain student thinking in ways which may be undesirable. An environment for beginners should require that the problem-solver bring the assumptions and methodology to the environment rather than the other way around.

Products of Phylogenetic Analysis

Researchers and students set out to explore a domain with different goals. Researchers seek to apply a particular set of principles to a body of data, while students seek to compare their own conceptions with a range of principles from across a domain. Phylogenetic representation in some ways is like a language: diagrammatic forms take on meaning when interpreted as representing biological phenomena. Research programs, like MacClade, embody a particular theoretical framework from within a research tradition. PI seeks to be neutral to research traditions so that student can explore how they compare with one another, where they overlap, and what

implications and assumptions each implies. PI permits students to play with a larger vocabulary in the language of phylogenetic representation.

MacClade produces cladistic hypotheses in a particular form. PI can produce cladistic hypotheses, but it can also be used to create phylogenetic hypotheses. Cladistic hypotheses show sister-group relationships among a group of taxa through the order of branching. Phylogenetic hypotheses show branching patterns of sister group relationship, but may also indicate ancestor/descendant relationships. Furthermore, in phylogenetic trees the location of taxa on the diagram, their proximity to each other, and the slope and length of links may represent theoretical hypotheses about the taxa, their relationships, and the underlying processes of evolution.

There is disagreement among scientists about the patterns that result from evolution and how those patterns should be represented. Cladists largely believe that fossil species should not be placed as direct ancestors to recent species while scientists from other research traditions hold that this practice is acceptable in the absence of contrary evidence (Harper, 1976). PI allows both forms of representation . Being able to represent both sister-group and ancestor/descendant relationships allows students to explore how these two forms of representation are related, what they mean, and when each is most appropriate.

Cladistic Hypotheses. Cladograms illustrate the nested groups that are expected as the result of evolution. Hennig and the early cladists idealized a particular form of diagram for representing the relationships created by the evolutionary process (Hennig, 1966; Maddison, 1989). This form has come to be called the Hennigian Comb. MacClade produces diagrams that approximate this format and has a tool (Ladderize) to modify any created diagram to, as closely as possible, assume this

form. Figure 1 shows a pair of MacClade diagrams that illustrate the ladderized form and an equivalent form. These diagrams represent the same data and mean exactly the same thing, from a cladistic standpoint. Cladograms only describe sister-group relationships. In Figure 1, Rfive is the sister group of Rfour, Rthree is the sister group to the monophyletic taxon of Rfive and Rfour, etc. Although evolutionary patterns are the result of ancestor/descendant relationship, cladists argue that phylogenetic analysis can only result in inferences about sister-group relationships. For that reason, all organisms whether recent or fossil, are placed along the top and lines that illustrate the hypothesized relationships are placed below. Location on the diagram conveys no information—only the pattern of branching is significant.

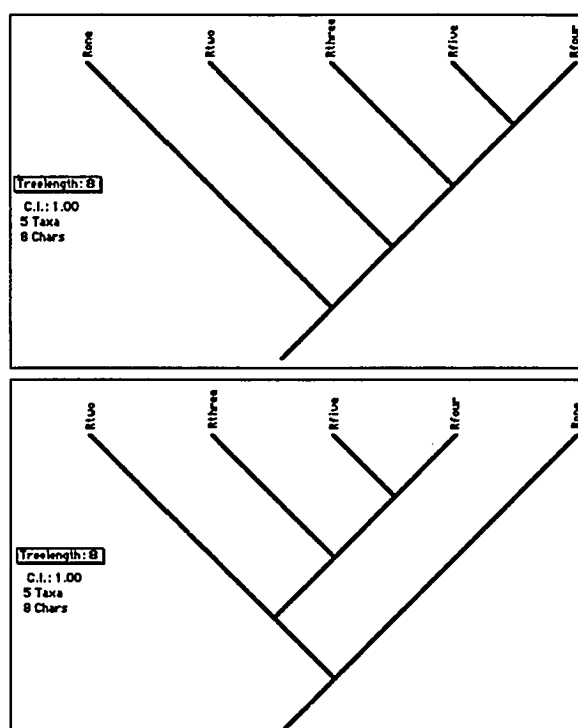


Figure 1. Equivalent Ladderized (Upper) and Non-Ladderized (Lower) Cladistic Representations Created Using MacClade 2.1.

Phylogenetic Investigator can create cladistic representations. Figure 2 illustrates how the hypothesis from Figure 1 can be represented cladistically using PI. There are differences between PI's cladistic diagrams and those produced using MacClade. These differences are primarily representational rather than conceptual in nature. PI's diagrams always contain a species at each branching point and display all of the character states transitions. Although character state transitions are not shown in MacClade, most cladists would assume their existence. By not showing them, however, MacClade does not stipulate their existence. One of the goals of PI is to present all of the hypothetical entities explicitly where students can evaluate their significance. This may also facilitate display-based reasoning on the part of students, as crucial features of the problem and its representation are never concealed from the problem-solver (Larkin, 1989).

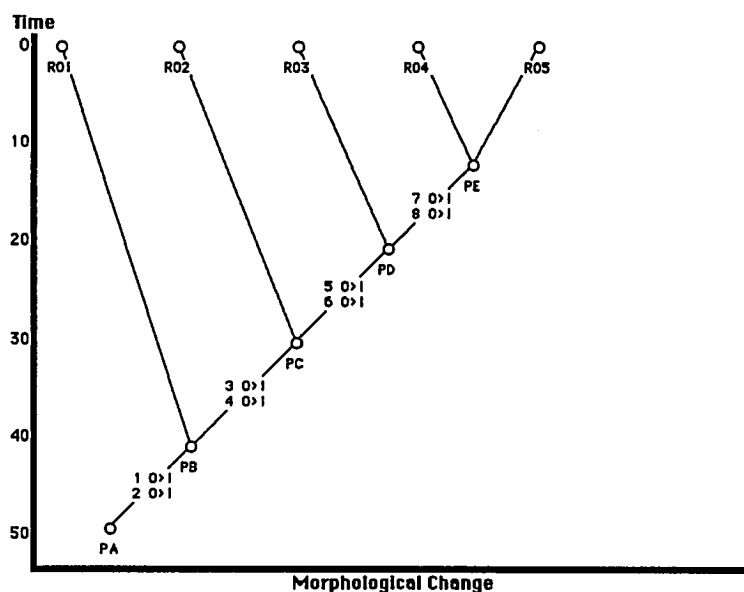


Figure 2. Cladistic Hypothesis Created Using Phylogenetic Investigator.

Phylogenetic Hypotheses. Phylogenetic trees can represent both sister-group and ancestor/descendant relationships. In Figure 3, the taxon F01 (a fossil taxon) is represented as the direct ancestor of both R01 and the monophyletic taxon of R02, F02, R03 and R04. In this way, students can construct hypothetical evolutionary pathways that illustrate how ancestor/descendant relationships could result in observed patterns of sister-group relationships. In addition, hypotheses can be represented that illustrate: taxa that appear more closely related; speciation events resulting from hybridization; variable rates of evolution and speciation; and points in time when particular evolutionary events are postulated to have occurred. These hypotheses are described here as "theoretical hypotheses."

Theoretical Hypotheses. Diagrams produced by MacClade are expected to represent the data in the data table and nothing more. Often, it is useful to modify representations to include theoretical hypotheses that can be used to explain

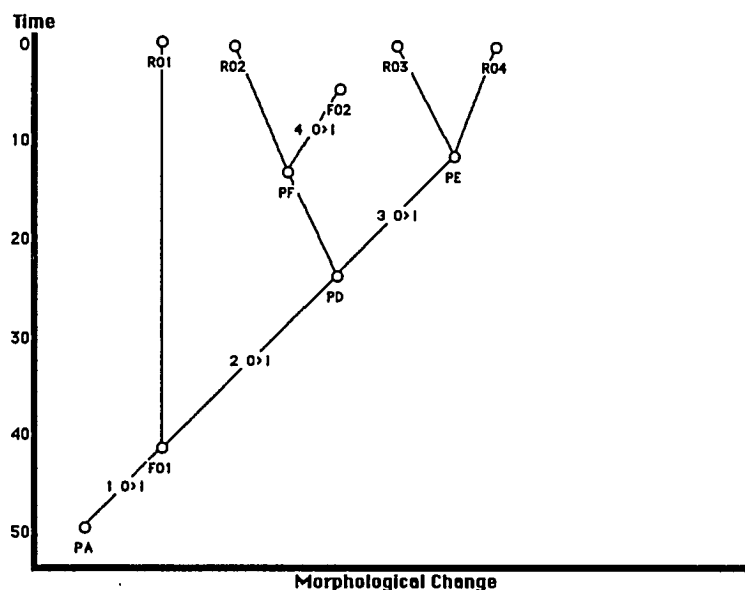


Figure 3. A Phylogenetic Hypothesis Incorporating Fossils as Ancestral (F01) and Sister (F02) Taxa.

differences observed between groups of organisms. These modifications can be added in artistically during the drawing process so that the diagram represents not only the data of phylogenetic analysis, but other more speculative observations which the student has made about the nature of the evolutionary pathways and processes operating in a particular problem. The observations and hypotheses created in this way can serve both as points of discussion during the persuasion process and as initial steps to guide students in gathering more data and extending their study of a group of organisms.

The rate of evolutionary change can be explored using PI by varying the angle of links between taxa (Figure 4). In some cases, the rate of change appears much greater between some members of a monophyletic taxon than others. This increased rate of change can be represented in PI by increasing the angle of a link at the point in the evolutionary history where the most change appears to have occurred. Figure 4 illustrates this theoretical hypothesis in the link between PD and PE. On the other hand, some taxa appear to change very little over long periods of time. Terminal taxa that show this characteristic are sometimes called "living fossils" and can be represented by linking them to an ancestor with a vertical line. Figures 3 and 4 each contain a taxon connected to an ancestor using a vertical line. In Figure 3, this ancestor is a fossil and in Figure 4, it is a postulated taxon.

Another form of variation in evolutionary rates is the time between speciation events. The theory of gradualism, that predicts speciation events to be more or less evenly spaced, has been challenged by a newer theory: punctuated equilibrium (Eldredge & Gould, 1972). Punctuated equilibrium proposes that speciation events may occur close together in time separated by long periods with relatively little change. PI can create hypotheses that illustrate punctuated equilibrium (Figure 5). In

Figure 5, all of the speciation events that give rise to the terminal taxa are hypothesized to have taken place between 24 and 26 units of time before the present.

Hybridization presents special problems to phylogenetic systematics (Funk, 1985) and is the most prominent form of a larger problem of reticulating genetic transfer. Representing hypotheses of complex ancestry has been similarly problematic, but most solutions involve the generation of a reticulated phylogenetic tree (Wagner, 1983). No current phylogenetic programs address the construction or analysis of reticulating trees (Haszprunar, 1992). PI can produce reticulate representations of phylogeny (Figure 6) and the issues surrounding this topic can be explored with students. Figure 6 represents taxon R03 as species that arose through past hybridization between the ancestors of taxa R01 and R02.

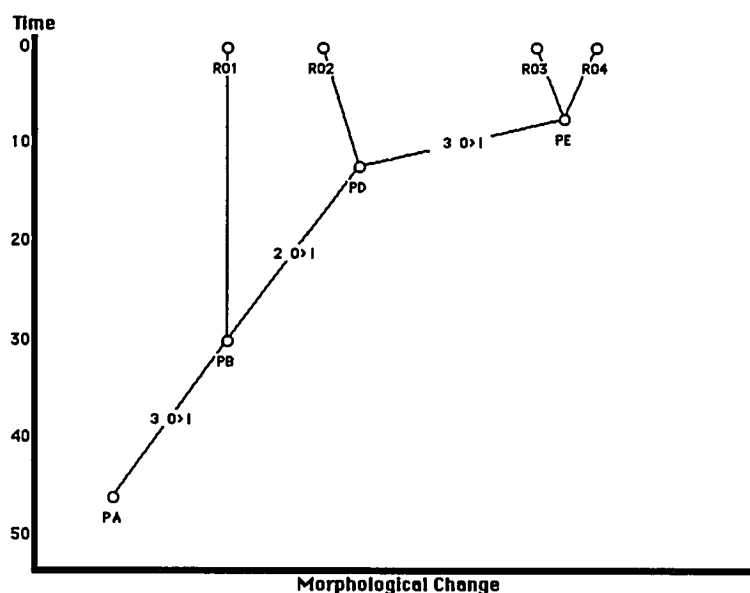


Figure 4. A Theoretical Hypothesis Showing R01 as a Living Fossil and an Increased Rate of Change Between PD and PE.

Other possible evolutionary patterns which might be addressed include: comparisons between the increasing cone of diversity vs. the model of diversification

and decimation (Gould, 1989), elastic diversification (Stanley, 1977), and comparing evolutionary vs. non-evolutionary systems (David Cowan, personal communication). Although I haven't provided the data that students might use to address these issues or support structuring their representations according to these theoretical frameworks, the capability to create them is important in order to explore the range of hypothesizing possible within phylogenetic analysis.

The ability of PI to encompass the full range of phylogenetic representation allows students to explore the limits of phylogenetic representation and learn what it means to exceed them. Students are capable of producing trees that contain impossible or meaningless relationships. But students often bring alternate conceptions to the classroom and should be able to find ways to represent these ideas and bring them out into the open where they can be discussed and remediated.

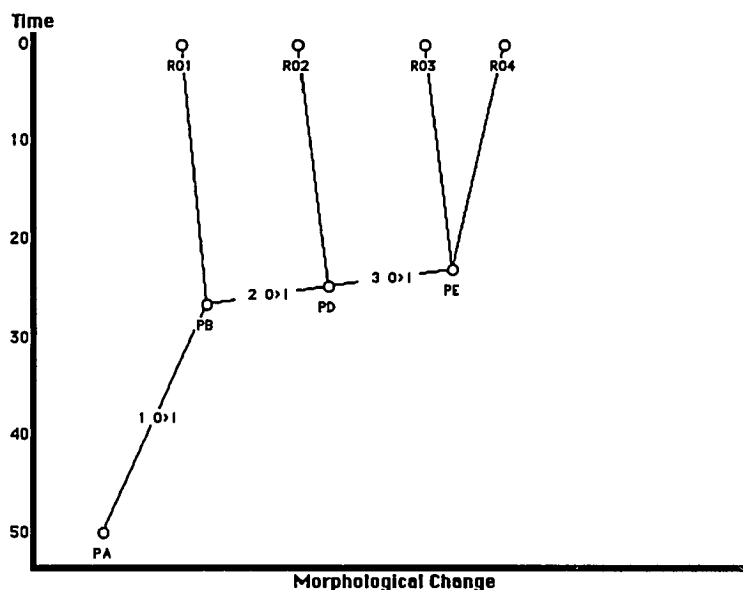


Figure 5. A Theoretical Hypothesis Showing Punctuated Equilibrium With One Period of Speciation.

Conclusions

The products and processes of problem solving are intertwined in complex ways. Students can use a flexible medium like PI to think about macroevolution in ways that a more constrained medium would not allow. The ability to encompass a wider variety of hypotheses gives students a greater opportunity to explore the possible forms that a representation of evolutionary relationships might take. With the few constraints in the PI medium, students are capable of producing trees that express their ideas about the evolutionary process. Second, the diagrams place all of the data where it can be readily seen and even allows students to represent qualitative judgments about the relationships represented. Finally, as students discuss their hypotheses, they have a greater opportunity to be reflective about the shortcomings of any particular representation and explore how else the data might be represented.

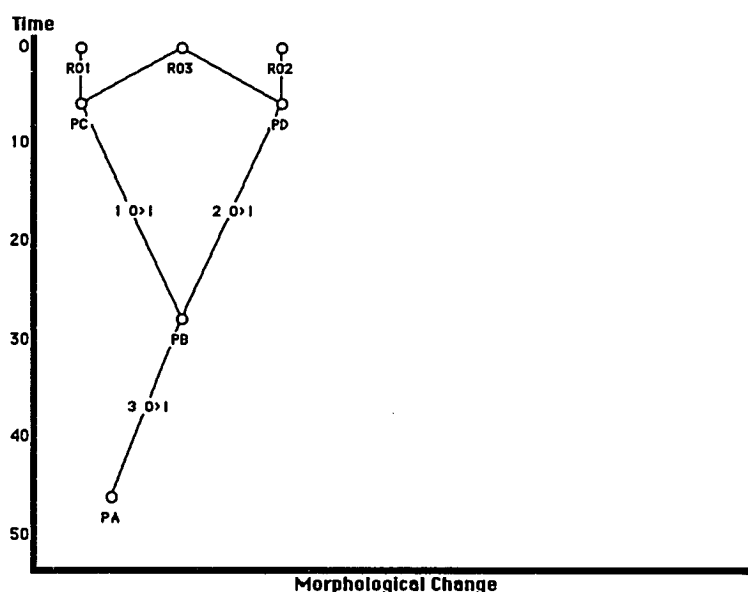


Figure 6. A Theoretical Hypothesis Showing Hybridization.

These are essential aspects of the persuasion component of the 3 P's (Stewart & Jungck, 1994).

Students that understand the assumptions of phylogenetic inference and the range of representational capability within the phylogenetic medium are well positioned to benefit from subsequent experience with research oriented tools. By providing an environment where students can explore these assumptions and have greater control over the resultant phylogenetic representations, PI offers the potential to improve how phylogenetic inference is taught to beginning students.

Rational Analysis

The goal of the rational analysis was to construct an initial process model that embodied the principles of phylogenetic inference that was sufficient to construct phylogenetic trees. As suggested by Larkin and Rainard (1984), this model was constructed in the form of condition and action couplets that identified the actions that certain conditions warranted. Table 4 brings together the principles of phylogenetic analysis into 6 "rules" structured as condition-action couplets that can be applied to problems of phylogenetic analysis to construct a solution. The term "rule" is applied here in a very general sense, meaning generalization or principle. I assumed experts had additional rules to recognize more complex phenomena and construct solutions.

The rules described here can be applied to at least simple problems of phylogenetic analysis to construct phylogenetic trees. Figure 7 shows a flow diagram that represents the steps necessary to create a tree for a phylogenetic problem using these rules. Each condition is illustrated with a representation of the data matrix and a box enclosing the relevant group of derived character states. The problem used for this example is described in greater detail below as Homoplasy 4. The rules of the

rational analysis served as the first process model used in the analysis of the problem-solving protocols.

Table 4
A Rational Analysis of Phylogenetic Tree Construction

Rule	Condition	Action
1.	If two taxa separated in time share identical states	then they can be linked.
2.	If there is (1) a largest group of taxa or (2) two mutually-exclusive groups of taxa defined by apomorphic character states	then a common ancestor or ancestors can be postulated supported by that character or characters.
3.	If a taxon from a group of taxa that share a character in the apomorphic state has already been identified with a postulated ancestor based on another character	then that character is considered uniquely apomorphic for that taxon.
4.	If a character is uniquely apomorphic for some taxa and shared by others	then a transition is placed at the most recent common ancestor to those that share it and the reversal of this character is uniquely apomorphic for those that do not share it.
5.	If a taxon possesses only uniquely apomorphic characters	then it can be linked to the most similar postulated ancestor and transitions for those uniquely apomorphic characters can be placed on that link.
6.	If there is a conflict between rules	then the rule which will result in the fewest transitions is selected unless there is an equal number of transitions in both cases when both rules will be accepted resulting in two equally parsimonious hypotheses.

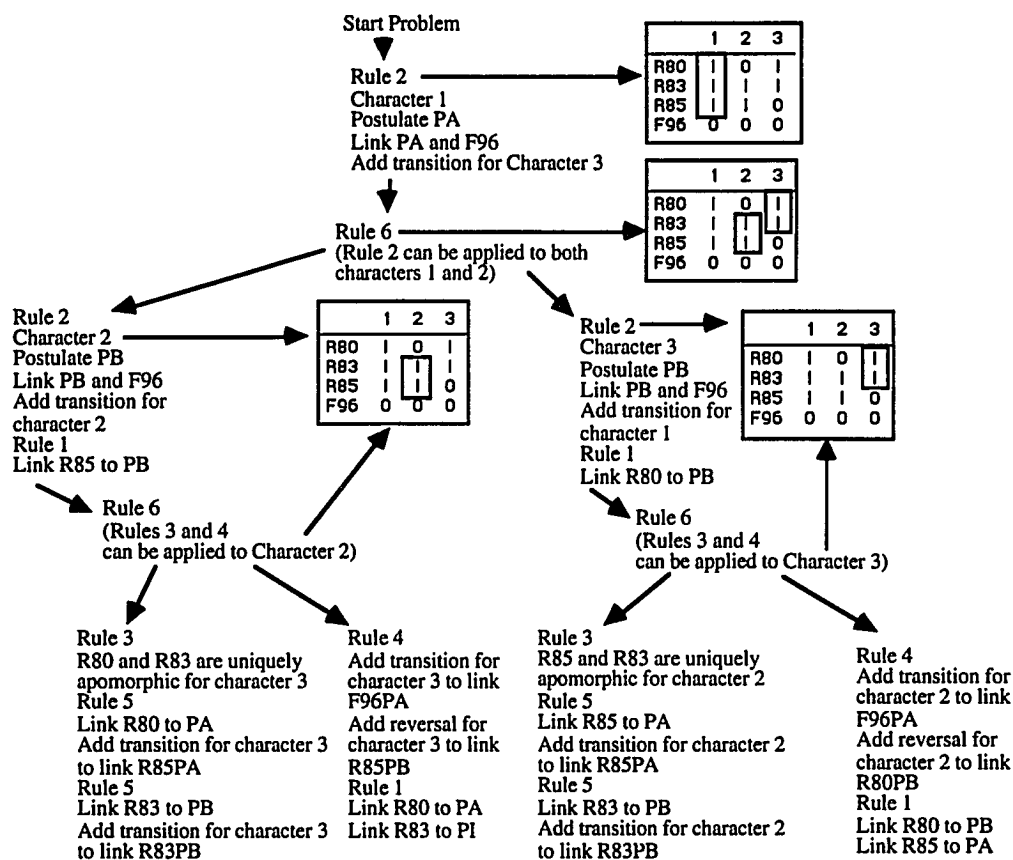


Figure 7. Applying Rational Analysis in Problem Solving.

Defining the Nature of the Problems

The nature of phylogenetic problems was considered from two standpoints. First, a series of model problems was constructed that represented the minimum amount of data to unambiguously define specific forms of solutions. The model problems were constructed to guide the search for organization of expert knowledge. Second, a series of research problems was constructed that attempted to survey problem complexity across numbers of solutions, characters, and taxa. The research problems were used for gathering data with experts.

A Typology of Model Phylogenetic Problems

The following problems represent all of the fundamental components from which more complex phylogenetic trees are constructed. Among the first five, the fundamental difference lies in the hierarchical scale of the results of a speciation event: Autapomorphies yield single species, synapomorphies result in a pair of species, a species and monophyletic taxon, or two monophyletic taxa. The two homoplasy problems represent tree topologies that result from reversal or convergence of characters. Each problem assumes Wagner parsimony.

The data sets associated with the following diagrams represent the minimum amount of data necessary to unambiguously define a particular tree topology or set of topologies. The Autapomorphy, Synapomorphy 1, Synapomorphy 2, Synapomorphy 3, and Homoplasy 3 problems have 1 topology and 1 optimization. The Homoplasy 1 & 2 problem has 1 topology with 2 character optimizations. The Homoplasy 4 problem has 2 topologies each with 2 character optimizations. Each problem has a set of recent taxa and a taxon that instantiates the ancestral condition.

The Autapomorphy problem (Figure 8) has 1 taxon, 1 character, and 1 solution (Figure 9) that requires 1 step. This problem demonstrates the essence of the phylogenetic problem: A taxon at one point in time (F97) is plesiomorphic (0) with respect to a character of interest (1) while a recent taxon (R85) has the character in the

	1
R85	1
F97	0

Figure 8. An Autapomorphy Data Matrix.

apomorphic state (1). The problem can be resolved by establishing a link of ancestral-descendant relationship and placing a transition for the character on the link.

The Synapomorphy 1 problem (Figure 10) has 2 taxa, 1 character, and 1 solution (Figure 11) that requires 1 step. This problem demonstrates the fundamental assumption of modern phylogenetic tree construction (or what is sometimes called "Hennig's auxiliary rule"). The two recent taxa (R81 and R83) share an apomorphic character which is plesiomorphic in the ancestor (F92). A common ancestor can be postulated, linked to the recent taxa and to the ancestor, and the transition for the character can be placed prior to the common ancestor.

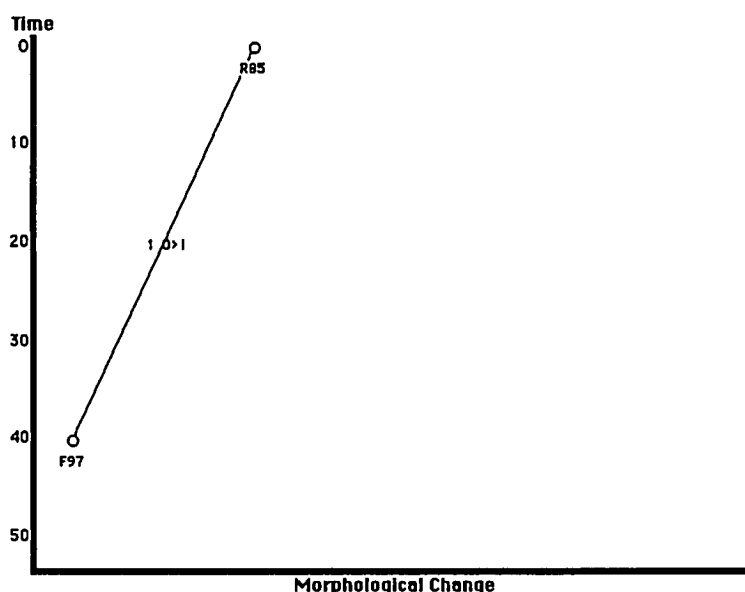


Figure 9. A Phylogenetic Tree Representing the Autapomorphy Problem.

	1
R81	1
R83	1
F92	0

Figure 10. A Synapomorphy 1 Data Matrix.

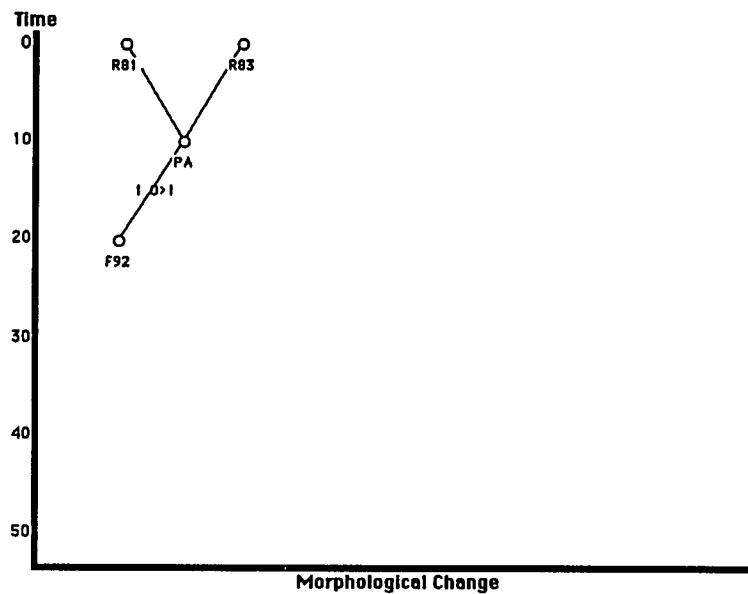


Figure 11. A Phylogenetic Tree Representing the Synapomorphy 1 Problem.

The Synapomorphy 2 problem (Figure 12) has 3 taxa, 2 characters, and 1 solution (Figure 13) that requires 2 steps . This problem illustrates nested characters (or the inclusion role): character 2's distribution of apomorphies (which defines a group of R84 and R85) is included entirely within character 1's distribution (which defines a group of R84, R85 and R89). Being inclusive is one way that characters can be "consistent" or "compatible." Inclusive characters represent a stronger hypothesis than exclusive characters.

	1	2
R85	1	1
R89	1	0
R84	1	1
F97	0	0

Figure 12. A Synapomorphy 2 Data Matrix.

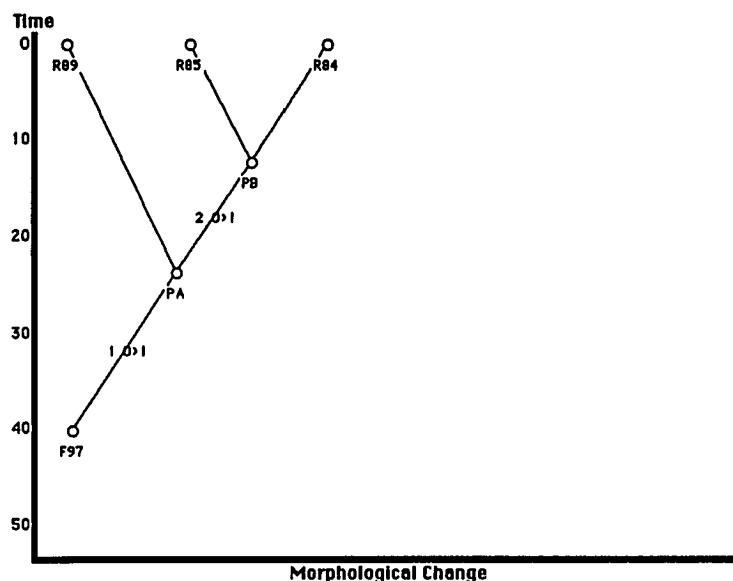


Figure 13. A Phylogenetic Tree Representing the Synapomorphy 2 Problem.

The Synapomorphy 3 problem (Figure 14) has 4 taxa, 3 characters, and 1 solution (Figure 15) that requires 3 steps. This problem illustrates mutually exclusive characters (or what is sometimes called "the exclusion rule"). Character 1 (that defines a group of R80 and R86) and character 2 (that defines a group of R82 and R84) are exclusive because their distributions do not overlap. Exclusive characters, like inclusive characters, are "compatible" or "consistent" with one another.

	1	2	3
R80	1	0	1
R82	0	1	1
R86	1	0	1
R84	0	1	1
F98	0	0	0

Figure 14. A Synapomorphy 3 Data Matrix.

The Homoplasy 1 & 2 problem (Figure 16) has 3 taxa, 4 characters, and a solution with 1 topology and 2 optimizations that requires 4 steps. This problem illustrates character optimizations. A character (1) conflicts with two identical characters (2 and 4) resulting in two different interpretations of the conflicting character. In the convergence interpretation (Figure 17), the conflicting character is gained twice (in R88 and R85). In the reversal interpretation (Figure 18), it is gained once (before PA) and lost once (in R80). (This problem is called Homoplasy 1 & 2 for historical reasons. During the development of the model problems, convergence and reversal were initially treated as separate phenomena. Use of the current designation avoids conflict with previous usage that would result from reorganizing the series.)

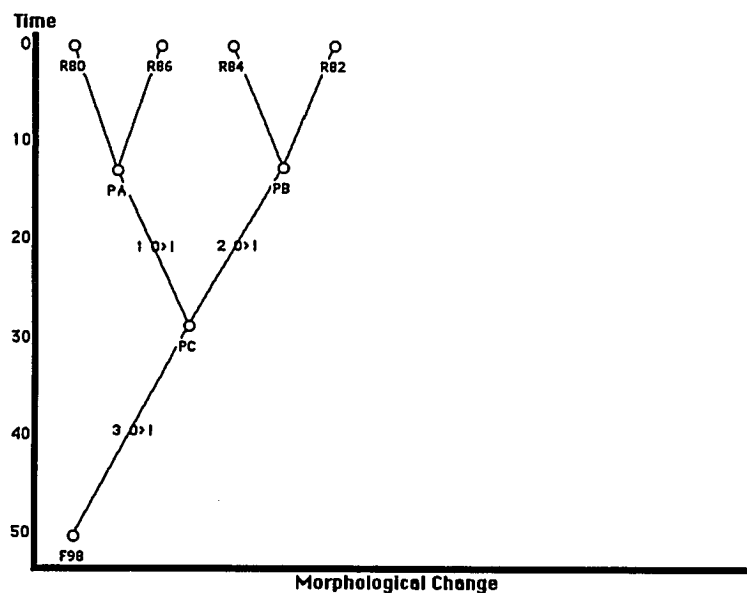


Figure 15. A Phylogenetic Tree Representing the Synapomorphy 3 Problem.

The Homoplasy 3 problem (Figure 19) has 4 taxa, 4 characters, 1 solution (Figure 20) that requires 5 steps. This problem illustrates homoplasy with a single

•

resolution (convergence). Two characters (1 and 3) are exclusive from each other and inclusive with respect to character 4. Character 2 conflicts with characters 1 and 3, but only one interpretation is possible in this case. Constructing this solution as a reversal, would be less parsimonious than the convergence optimization (3 steps would be required to explain the homoplasious character).

	1	2	3	4
R88	1	1	1	1
R80	0	1	1	1
R85	1	0	1	0
F99	0	0	0	0

Figure 16. A Homoplasly 1&2 Data Matrix.

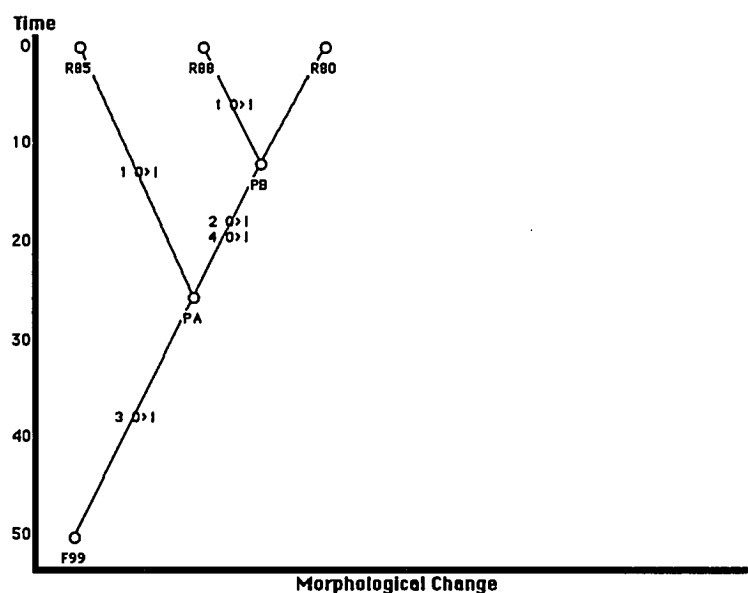


Figure 17. The Convergence Character Optimization for the Homoplasly 1&2 Problem.

The Homoplasy 4 problem (Figure 21) has 3 taxa, 3 characters, and a solution with 2 topologies, that each have 2 optimizations, that requires 4 steps. This problem illustrates multiple topologies. It is similar to the Homoplasy 1 & 2 problem, except

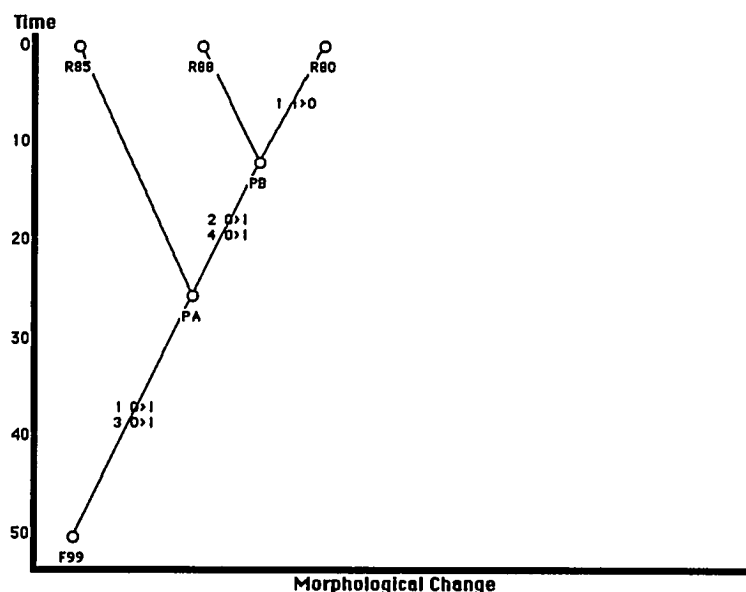


Figure 18. The Reversal Character Optimization for the Homoplasy 1&2 Problem.

	1	2	3	4
R84	1	1	0	1
R87	0	1	1	1
R80	1	0	0	1
R88	0	0	1	1
F94	0	0	0	0

Figure 19. A Homoplasy 3 Data Matrix.

that there are now only two characters (2 and 3) that conflict with each other. In the Homoplasy 1 & 2, the two identical characters unambiguously define the tree's structure. In this case, either character is "equally believable" resulting in two arrangements of the taxa, each with two character interpretations. In solution 1,

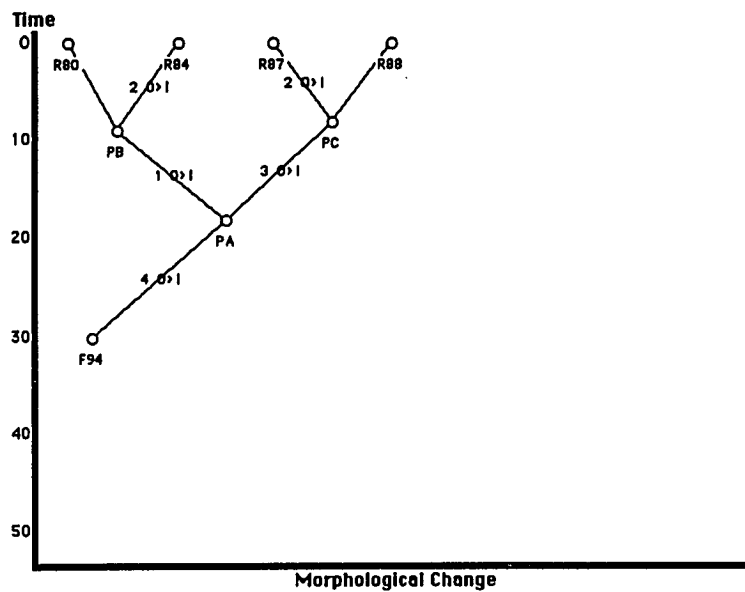


Figure 20. A Phylogenetic Tree Representing the Homoplasy 3 Problem.

	1	2	3
R80	1	0	1
R83	1	1	1
R85	1	1	0
F96	0	0	0

Figure 21. A Homoplasy 4 Data Matrix.

character 2 puts 83 and 85 together (Figure 22) and in solution 2, character 3 puts 80 and 83 together (Figure 23). Each topology also has 2 character optimizations for the homoplasious character that can be represented as either a convergence or a reversal.

Research Problems

For the purposes of this study, a problem is defined as (a) a task in which a solver (b) explores a specific environment using permissible operations to (c) attain a specified goal from (d) a specified initial state (Reif, 1983; Collins, 1986). In this

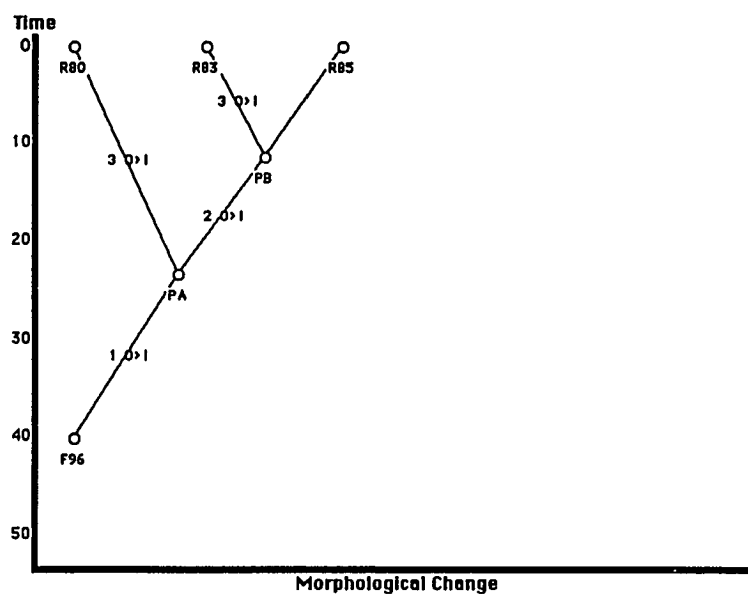


Figure 22. Solution 1 of Two Equally Parsimonious Solutions for the Homoplasy 4 Problem.

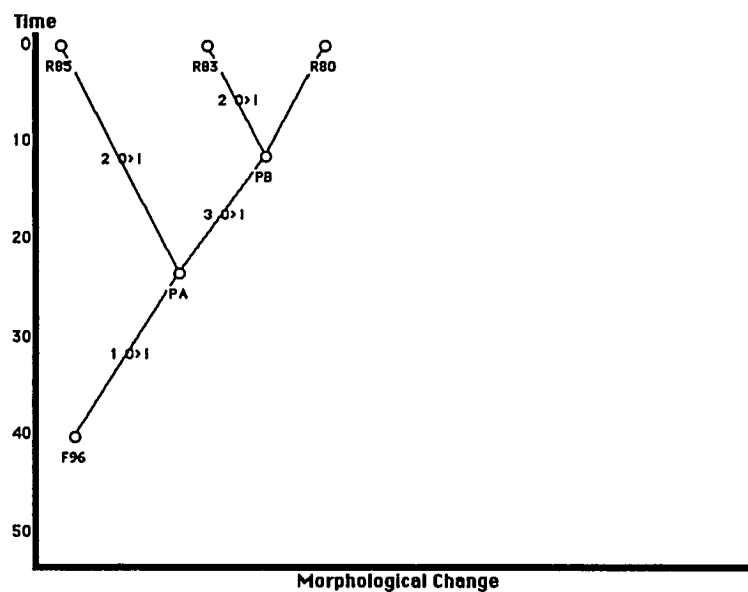


Figure 23. Solution 2 of Two Equally Parsimonious Solutions for the Homoplasy 4 Problem.

study a problem is: (a) the task of phylogenetic tree construction; (b) using permissible operations governed by the problem-solving environment provided by Phylogenetic Investigator or the representational capabilities of the participant using paper and pencil; (c) to attain the goal of construction of one or more most-parsimonious phylogenetic trees for each problem; and (d) from the specified initial state of the set of taxa, characters, and states presented in each research problem.

In constructing the set of research problems, I sought to use the primary factors that govern complexity to elicit as many strategies and heuristics as possible. In addition, I created problems to examine the revision of existing trees under a variety of circumstances and a series where fossil evidence was available in a variety of different relationships with the ingroup. The number of the series represents the order in which the set of problems was constructed during the development of the problems

Complexity was viewed as related to numbers of solutions, characters, and taxa. Table 5 presents a summary of the research problems. Series 1 was composed of problems with 5 taxa, 5 characters, and different numbers of solutions. Series 4 was composed of problems with 5 taxa and a single solution, but with different numbers of characters (10, 15, or 20). Series 5 was composed of problems with 10 characters and a single solution, but different numbers of taxa (6, 8 and 10).

Each problem in Series 2 dealt with revision under a variety of circumstances. In each, the participant would solve a problem that had 5 characters and 5 taxa and then reexamine the problem with additional data. Additional data could take the form of a taxon (in problems 2.1, 2.3, and 2.4) or 2 characters (in problem 2.2). The revisions required included simple addition of taxa to a tree, significant restructuring of an existing tree, and increases or decreases in the number of valid solutions.

Table 5
Summary of Research Problems

Series 1 (Ambiguity axis): 5 taxa, 5 characters held constant

Problem	Variables	
1.1	1 solution	1 optimization
1.2	1 solution	2 optimizations
1.3	2 solutions	2 optimizations
1.4	3 solutions	2 optimizations
1.5	4 solutions	2 optimizations

Series 2 (Revision problems): 5 taxa, 5 characters held constant

Problem	Variables
2.1	1 new taxon that "fits" with previous taxa
2.2	2 new characters result in restructured tree
2.3	1 new taxon results in more valid solutions
2.4	1 new taxon results less valid solutions

Series 4 (Character axis): 5 taxa, 1 solution held constant

Problem	Variables
4.1	10 characters
4.2	15 characters
4.3	20 characters

Table 5—Continued

Series 5 (Taxon axis): 10 characters, 1 solution held constant

Problem	Variables
5.1	6 taxa
5.2	8 taxa
5.3	10 taxa

An effort was made to ensure that the problems also contained all of the fundamental problem types identified in the typology of problems. Autapomorphies appear in 2.2 and 2.4 Synapomorphy Problems 1, 2, and 3 appear in research problem 1.1. A Homoplasy 1 & 2 type problem appears in 1.2. A Homoplasy 3 type problem appears in research problem 2.2. A Homoplasy 4 type problem appears in research problem 1.3.

A set of imaginary taxa (R80-R89) were used to construct each problem. Each problem consisted of a set of polarized binary characters for a randomly selected group of taxa combined with an outgroup taxon F98, which was always composed solely of zeros. The problems were constructed using a word processor and then tested using PAUP (Swofford, 1991) to ensure that they possessed the desired characteristics. Each unique topology contained within a solution for a problem is identified by a solution number. These numbers are arbitrary for all purposes here and reflect the order in which topologies were reported by PAUP.

There was no clear precedent to guide how complex problems would need to be in order to challenge experts nor was it clear how many problems an expert would be able to solve before becoming fatigued. I therefore constructed more problems

than I thought could likely be solved, tried to bracket complexity on each axis, and then used the performance of the first participant to guide reducing the problem-set to a manageable body. Time and fatigue proved to be a considerable factor. Series 3 (dealing with use of fossil evidence) was dropped and series 4 and 5 (complexity across numbers of characters and taxa) were reduced to single problems. In those cases the most complex examples were retained, as S1 had demonstrated that they were solvable, though challenging.

Problem 1.1 (Figure 24) has a whole-group synapomorphy defined by character 1. Character 5 is nested within two identical characters, 3 and 4, which are exclusive from character 2. This results in a single tree with no homoplasy. The single most parsimonious solution (Figure 25) requires 5 steps.

SPC	1	2	3	4	5
R80	1	1	0	0	0
R89	1	0	1	1	0
R81	1	0	1	1	1
R82	1	1	0	0	0
R86	1	0	1	1	1
F98	0	0	0	0	0

Figure 24. The Data Matrix for Problem 1.1.

Problem 1.2 (Figure 26) has no evident whole-group synapomorphy. Character 4 is nested within identical characters 1 and 2, exclusive from 5. Character 3 requires homoplasy. Character 3 can be constructed either as two multiple gains (Figure 27) or be postulated as a synapomorphy for the whole group with a loss in common ancestor of 84 and 86 (Figure 28). Both arrangements represent a single most parsimonious solution that requires 6 steps.

Problem 1.3 (Figure 29) has no evident whole ingroup synapomorphy. Characters 3 and 5 are identical and exclusive from 2. Characters 1 and 4 conflict with each other, resulting in two equally parsimonious solutions. In solution 1 (Figure 30), character 1 defines a group of R85 and R87 and in solution 2 (Figure 31) character 4 defines a group of R85 and R84. For each solution, the conflicting character can be optimized as either two gains or a gain and a loss. The 2 most parsimonious solutions require 6 steps.

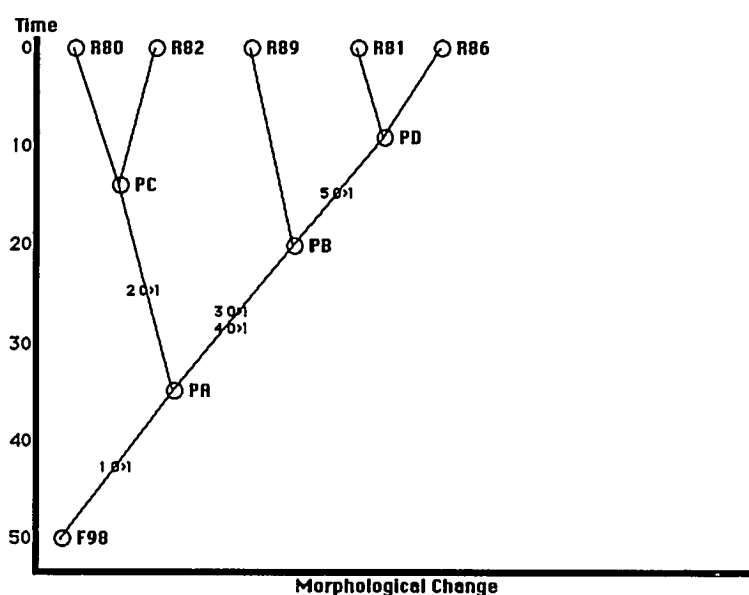


Figure 25. The Most Parsimonious Solution for Problem 1.1.

SPC	1	2	3	4	5
R89	0	0	1	0	1
R81	1	1	1	0	0
R83	0	0	1	0	1
R84	1	1	0	1	0
R86	1	1	0	1	0
F98	0	0	0	0	0

Figure 26. The Data Matrix for Problem 1.2.

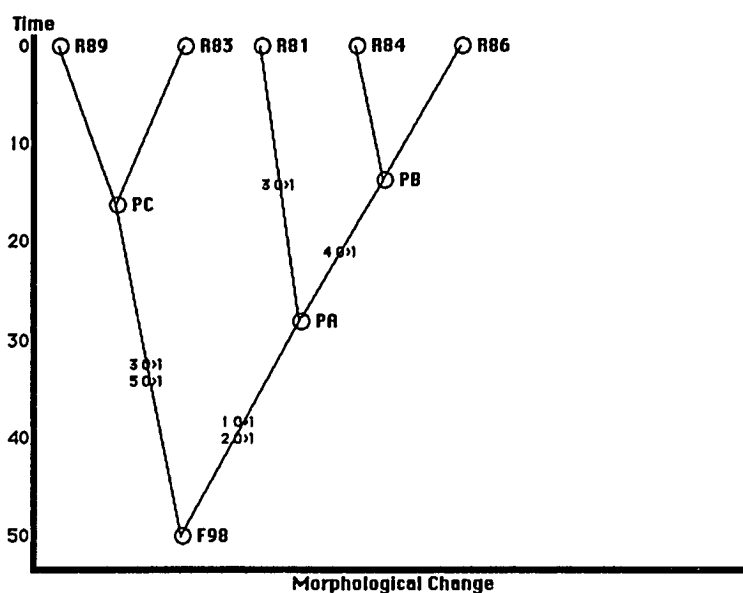


Figure 27. One of Two Equally Parsimonious Character Optimizations for Problem 1.2: The Convergence Option.

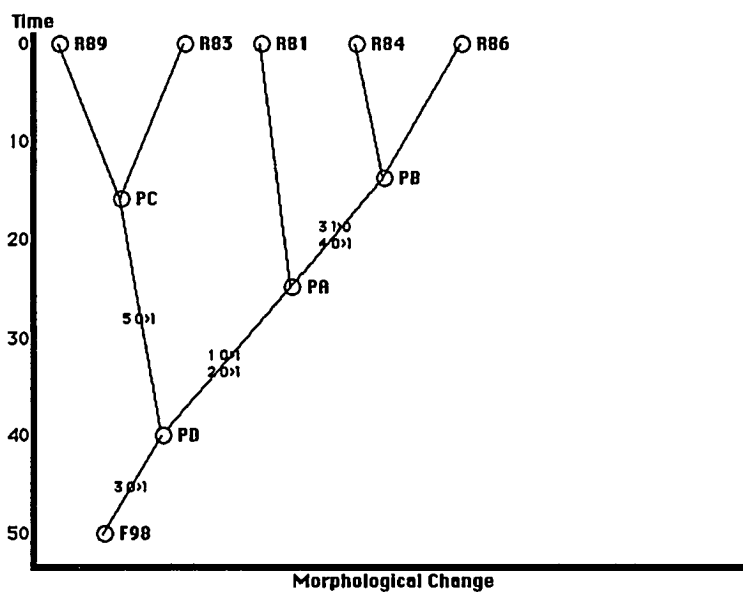


Figure 28. One of Two Equally Parsimonious Character Optimizations for Problem 1.2: The Reversal Option.

SPC	1	2	3	4	5
R83	0	1	0	0	0
R81	0	1	0	0	0
R87	1	0	1	0	1
R84	0	0	1	1	1
R85	1	0	1	1	1
F98	0	0	0	0	0

Figure 29. The Data Matrix for Problem 1.3.

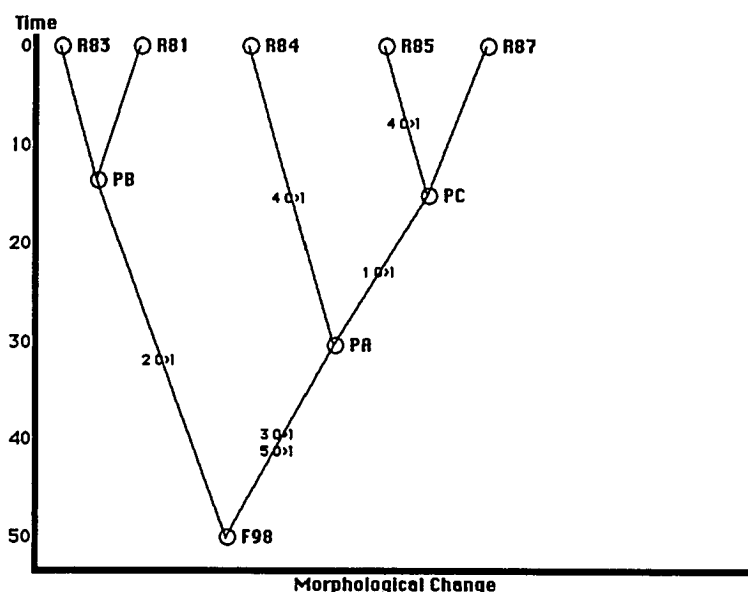


Figure 30. Solution 1 of Two Equally Parsimonious Solutions for Problem 1.3.

Problem 1.4 (Figure 32) has no evident whole-ingroup synapomorphy.

Characters 3 and 5 are identical and exclusive from character 2. Character 1 conflicts with character 4. Character 4 conflicts with all other characters. Three topologies are possible: Solution 1 (Figure 33) involves homoplasy in 4 and 1 (4 gained in R83 and the common ancestor of R89 and R81, and 1 either gained in R86 and R81 or gained in the common ancestor of R86, R89 and R81 and lost in R89); solution 2 (Figure 34) involves homoplasy in characters 2 and 4 (character 2 gained in R80 and R83,

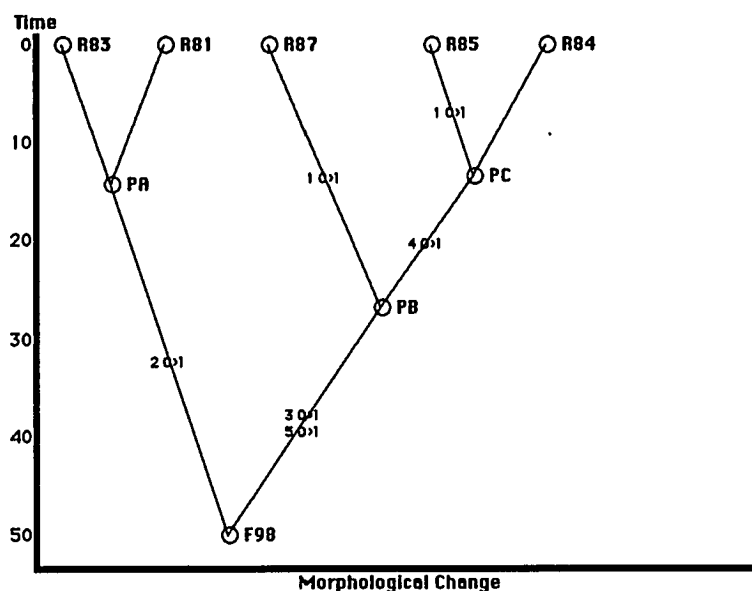


Figure 31. Solution 2 of Two Equally Parsimonious Solutions for Problem 1.3.

SPC	1	2	3	4	5
R83	0	1	0	1	0
R80	0	1	0	0	0
R86	1	0	1	0	1
R89	0	0	1	1	1
R81	1	0	1	1	1
F98	0	0	0	0	0

Figure 32. The Data Matrix for Problem 1.4.

character gained in the common ancestor of R83, R89, R81, and R86 and then lost in R86); and solution 3 (Figure 35) involves homoplasy only in character 4, which requires 3 gains (in R83, R89, and R81) or two gains and a loss (gained prior to the common ancestor of R89, R86, and R81 and lost in R86). The 3 most parsimonious solutions require 7 steps.

Problem 1.5 (Figure 36) has no evident whole-ingroup synapomorphy. Characters 3 and 5 are identical and exclusive from 2. Characters 1 and 4 both

conflict with each other and with all other characters. There are 4 most parsimonious solutions. Two solutions assume character 2 to be true and the other two require

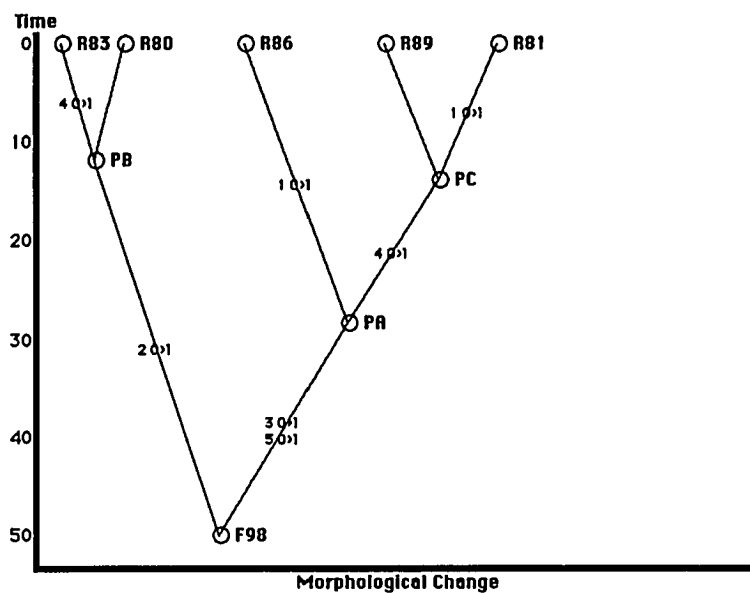


Figure 33. Solution 1 of Three Equally Parsimonious Solutions for Problem 1.4.

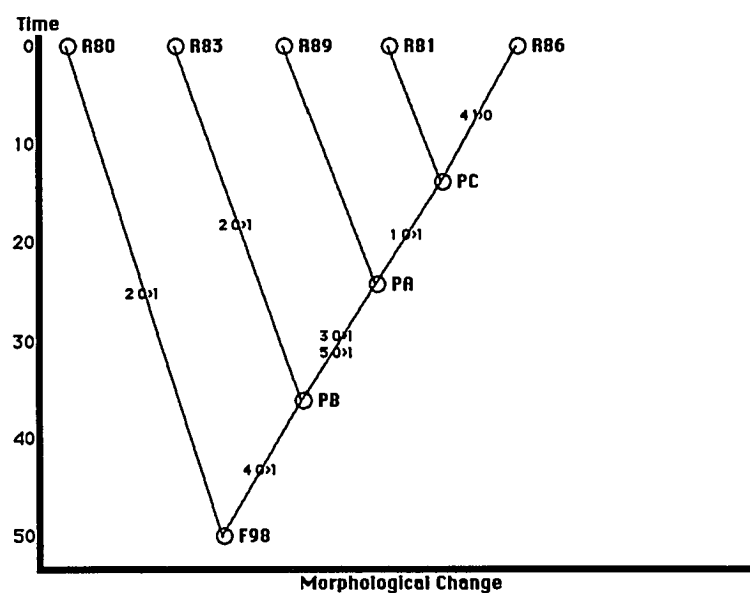


Figure 34. Solution 2 of Three Equally Parsimonious Solutions for Problem 1.4.

homoplasy in character 2. Rearrangements among taxa within R82, R88 and R83 under each condition of character 2 result in the other topologies. In the first of these solutions (Figure 37), character 4 is used to group R88 and R83 with character 1 either appearing as two gains (in R88 and R82) or as a gain in the common ancestor of R83, R88, and R82 and then a loss in R83. In the second solution (Figure 38),

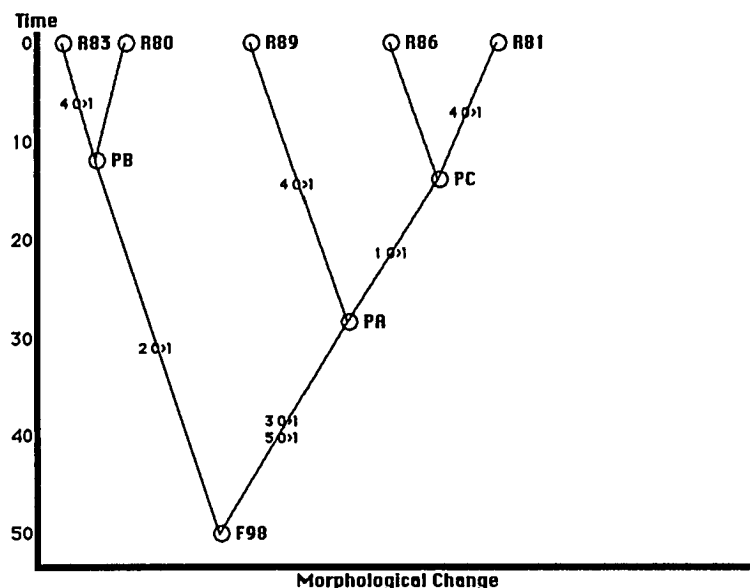


Figure 35. Solution 3 of Three Equally Parsimonious Solutions for Problem 1.4.

character 4 is hypothesized to transform after the divergence of R84 and then to reverse in R82. In the third solution (Figure 39), character 2 groups 84 and 86, character 1 is used to group R88 and R82 with character 4 either appearing as two gains (in R88 and R83) or as a gain in the common ancestor of R83, R88, and R82 and then a loss in R82. In the fourth solution (Figure 40), character 1 is hypothesized to transform after the divergence of R86 and then to reverse in R83. The 4 most parsimonious solutions require 8 steps.

SPC	1	2	3	4	5
R82	1	0	1	0	1
R86	0	1	0	1	0
R84	1	1	0	0	0
R88	1	0	1	1	1
R83	0	0	1	1	1
F98	0	0	0	0	0

Figure 36. The Data Matrix for Problem 1.5.

Problem 2.1 (Figure 41) has no evident whole-ingroup synapomorphy. Characters 2 and 4 are identical and nested within character 5 which is exclusive from 1. Character 3 is gained in the common ancestor of R85 and R86 and homoplasiously in R88. The single most parsimonious solution (Figure 42) requires 6 steps. Upon revision (Figure 43), a taxon, R87, has been added which has apomorphies for characters 2 and 5, but not for character 4. This taxon fits between R89 and R82

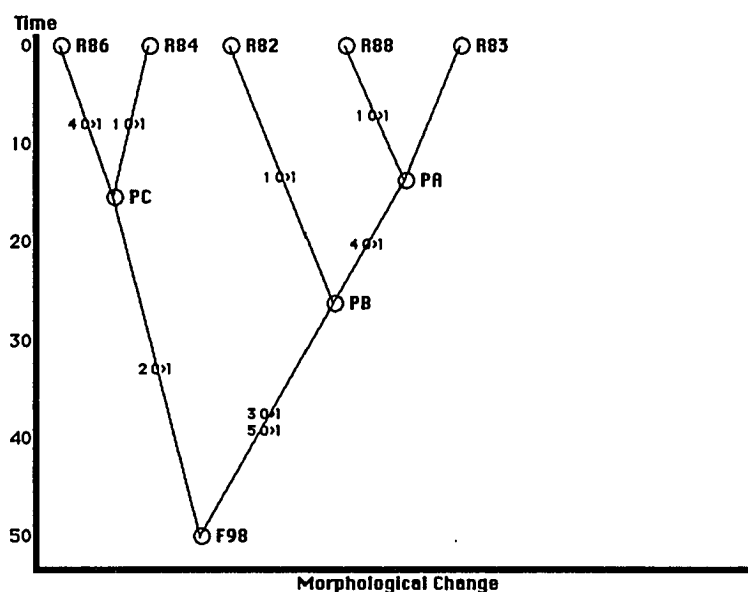


Figure 37. Solution 1 of Four Equally Parsimonious Solutions for Problem 1.5.

(Figure 44) and fixes the order in which characters 2 and 4 were transformed. The single most parsimonious solution still requires only 6 steps.

Problem 2.2 (Figure 45) has no evident whole ingroup synapomorphy. Character 4 is an autapomorphy. Character 3 defines a group of 4 taxa with characters 1 and 5 nested within 3 and exclusive from each other. Character 2 must be gained homoplasiously in R81, R80, and R88. The single most parsimonious solution (Figure 46) requires 7 steps. Upon revision (Figure 47), this problem contains 2 additional

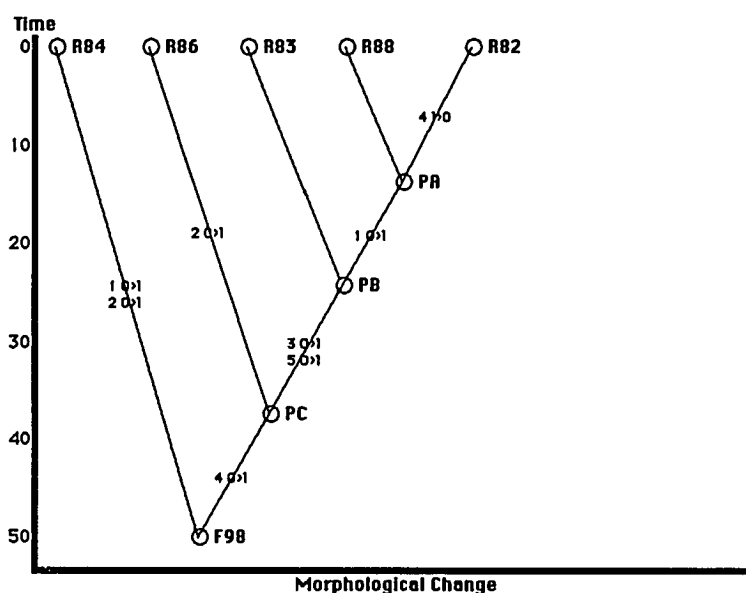


Figure 38. Solution 2 of Four Equally Parsimonious Solutions for Problem 1.5.

characters: character 6 groups R84, R81, and R88; and character 7 groups R81 and R88. These two characters result in the placement of R88 as a sister group to R81 with characters 1, 2, and 3 being homoplasious (Character 1 is gained in the common ancestor of R81, R88, and R84 and then lost in R88. Character 2 gained in the common ancestor of R81 and R88 and also in R80. Character 3 becomes a whole

ingroup synapomorphy which is lost in R88.) The single most parsimonious solution (Figure 48) requires 10 steps.

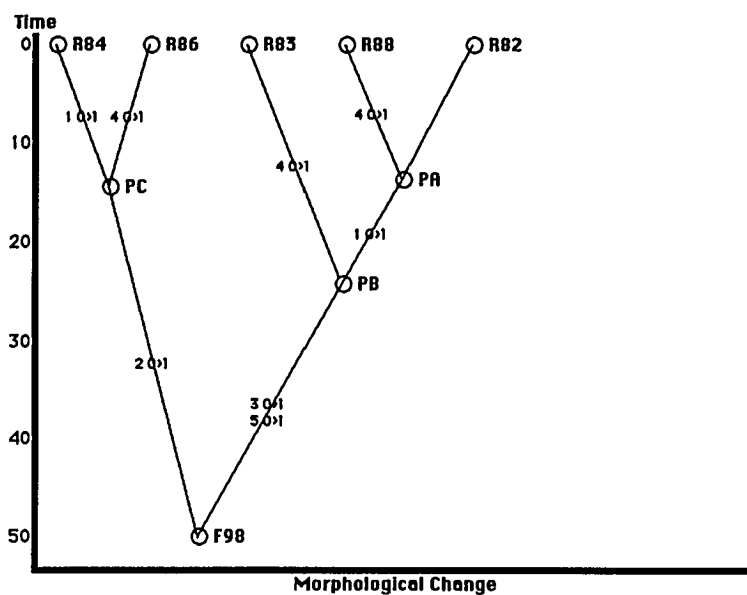


Figure 39. Solution 3 of Four Equally Parsimonious Solutions for Problem 1.5.

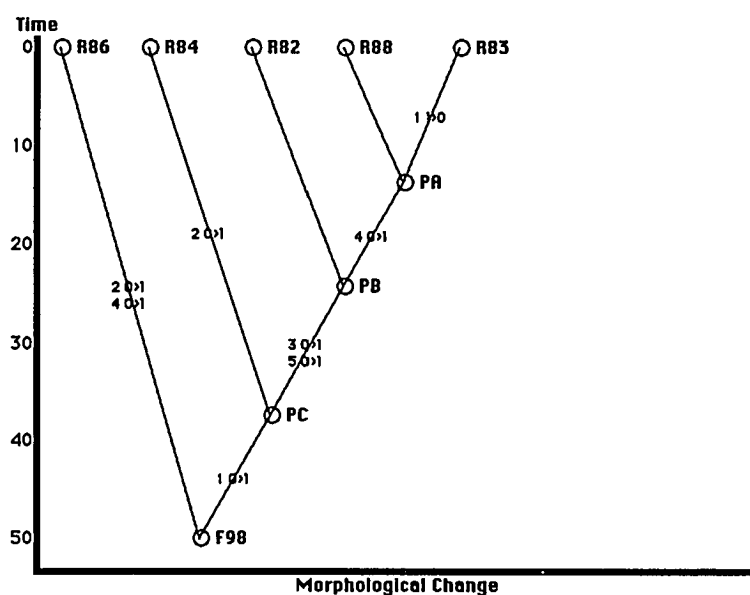


Figure 40. Solution 4 of Four Equally Parsimonious Solutions for Problem 1.5.

SPC	1	2	3	4	5
R88	0	1	1	1	1
R89	0	0	0	0	1
R85	1	0	1	0	0
R82	0	1	0	1	1
R86	1	0	1	0	0
F98	0	0	0	0	0

Figure 41. The Data Matrix for Problem 2.1.

Problem 2.3 (Figure 49) has no evident whole ingroup synapomorphy. Character 5 defines the largest group of taxa and includes characters 2 and 4, which are identical and character 3, which is exclusive from 2 and 4. Character 1 requires homoplasy and must be gained in R88 and R80. The single most parsimonious solution (Figure 50) requires 6 steps. Upon revision (Figure 51), this problem contains an additional taxon (R89) which has apomorphies for characters 1 and 4.

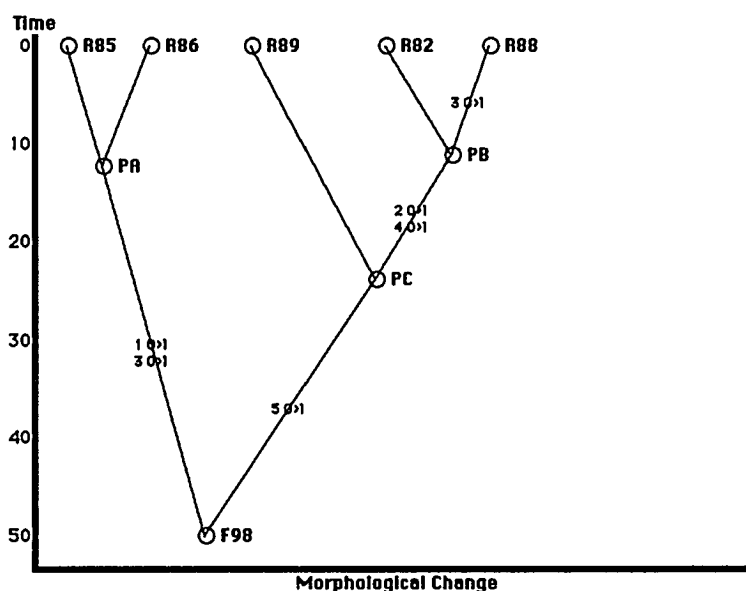


Figure 42. The Most Parsimonious Solution for Problem 2.1.

SPC	1	2	3	4	5
R88	0	1	1	1	1
R89	0	0	0	0	1
R85	1	0	1	0	0
R82	0	1	0	1	1
R86	1	0	1	0	0
R87	0	1	0	0	1
F98	0	0	0	0	0

Figure 43. The Data Matrix for Problem 2.1a.

This results in two most parsimonious solutions. In the first (Figure 52), character 4 groups R89, R88 and R86; character 5 appears homoplasiously prior to the common ancestors of R84, R81 and R80, R86; and character 1 appears in the common ancestor of R80, R89, R88, and R86 and is then lost in R86. The second solution (Figure 53) has R89 appearing on the same branch as R80 and homoplasiously acquiring character 4. The two most parsimonious solutions require 7 steps.

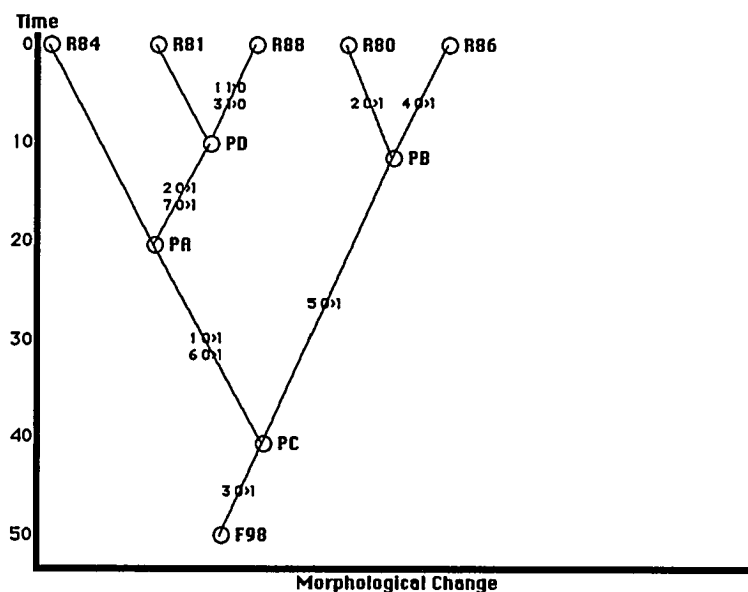


Figure 44. The Most Parsimonious Solution for Problem 2.1a.

SPC	1	2	3	4	5
R81	1	1	1	0	0
R80	0	1	1	0	1
R88	0	1	0	0	0
R86	0	0	1	1	1
R84	1	0	1	0	0
F98	0	0	0	0	0

Figure 45. The Data Matrix for Problem 2.2.

Problem 2.4 (Figure 54) has no evident whole ingroup synapomorphy. Character 3 is an autapomorphy. There are two most parsimonious solutions one based on characters 1 and 2 (Figure 55) and the other on characters 4 and 5 (Figure 56). If 1 and 2 are true, 4 must be homoplasiously gained in R81 and R86 and character 5 must be gained in the common ancestor of R89 and R86 and also in R81. If 4 and 5 are true, character 1 must be gained in the common ancestor of R82, R89,

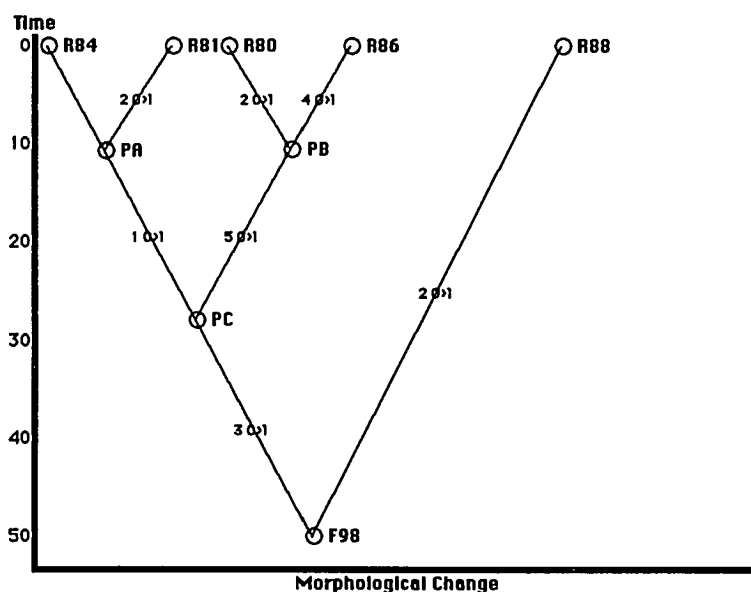


Figure 46. The Most Parsimonious Solution for Problem 2.2.

SPC	1	2	3	4	5	6	7
R81	1	1	1	0	0	1	1
R80	0	1	1	0	1	0	0
R88	0	1	0	0	0	1	1
R86	0	0	1	1	1	0	0
R84	1	0	1	0	0	1	0
F98	0	0	0	0	0	0	0

Figure 47. The Data Matrix for Problem 2.2a.

R81, and R86 and then lost in R86 and character 2 must be gained in R84 and R81. The two most parsimonious solutions require 7 steps. Upon revision (Figure 57), a taxon (R88) has been added with apomorphies for all the characters. This results in a single most parsimonious solution (Figure 58) where characters 4 and 5 are true. Character 1 must be gained in the common ancestor of R82, R89, R86, R88 and R81 and then lost in R86. Character 2 must be gained in R84 and R81. Character 3 must be gained in R89 and R88. The single most parsimonious solution requires 8 steps.

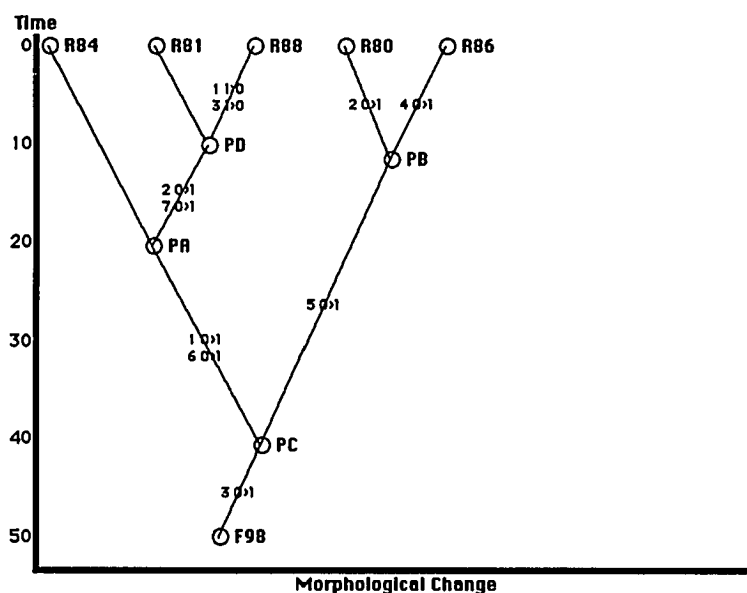


Figure 48. The Most Parsimonious Solution for Problem 2.2a.

SPC	1	2	3	4	5
R84	0	0	1	0	1
R86	0	1	0	1	1
R81	0	0	1	0	1
R80	1	0	0	0	0
R88	1	1	0	1	1
F98	0	0	0	0	0

Figure 49. The Data Matrix for Problem 2.3.

Problem 4.3 (Figure 59) has a whole ingroup synapomorphy (character 8). Characters 6 and 13 are identical and exclusive from 3 and 17. Characters 6 and 13 define R89, R80 and R82 with characters 1 (with Homoplasy in R81), 9 (with homoplasy in R85) and 10, defining R80 and R82 as sister taxa. All other characters

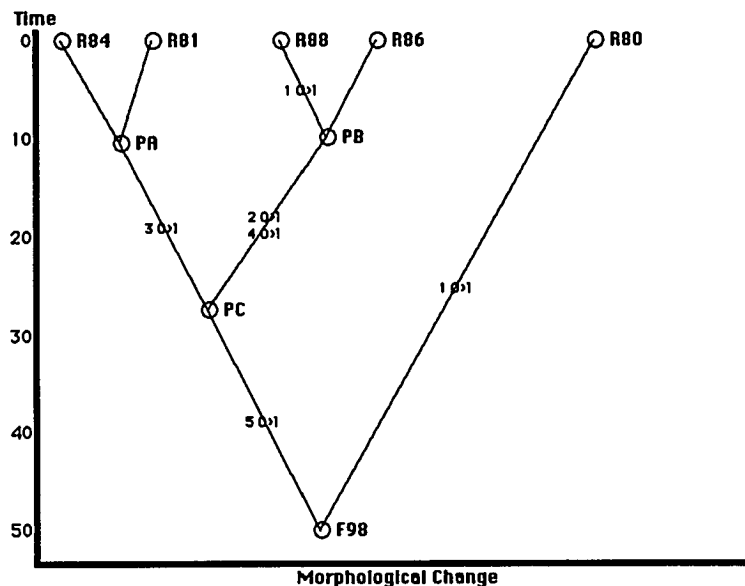


Figure 50. The Most Parsimonious Solution for Problem 2.3.

are homoplasious (Character 2 can be either gained in the common ancestor of R81 and R85 and also in R89 or gained in the common ancestor of the whole ingroup and

lost in the common ancestor of R82 and R89. Characters 4 and 16 can be gained in R89 and R80 or gained in common ancestor of R82 and R89 and lost in R82. Characters 5 and 15 must be gained in R85 and R89. Character 7 can be gained in R89 and R82 or gained in common ancestor of R82 and R89 and lost in R80.

SPC	1	2	3	4	5
R84	0	0	1	0	1
R86	0	1	0	1	1
R81	0	0	1	0	1
R80	1	0	0	0	0
R88	1	1	0	1	1
R89	1	0	0	1	0
F98	0	0	0	0	0

Figure 51. The Data Matrix for Problem 2.3a.

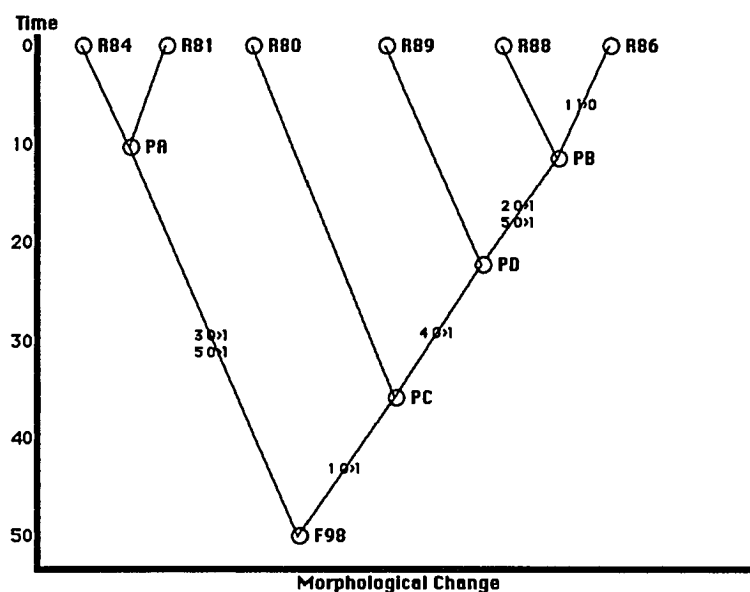


Figure 52. Solution 1 of Two Equally Parsimonious Solutions for Problem 2.3a.

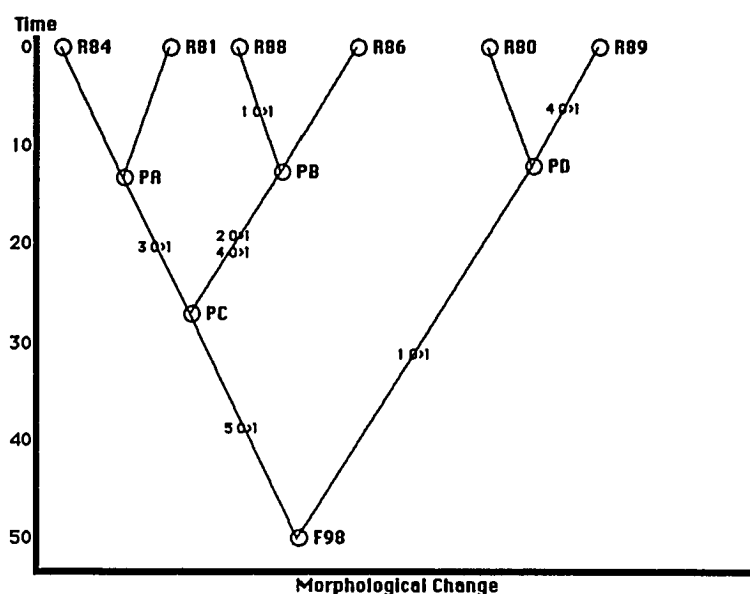


Figure 53. Solution 2 of Two Equally Parsimonious Solutions for Problem 2.3a.

Character 11 must be gained in R82 and R81. Character 12 must be gained in R81 and R89. Character 14 can be gained in R89 and R82 or gained in common ancestor of R82 and R89 and lost in R80. Character 18 must be gained in R81 and R80. Character 19 must be gained in R82 and R85. Character 20 must be gained in R80 and R85.) The single most parsimonious solution (Figure 60) requires 34 steps.

SPC	1	2	3	4	5
R86	1	0	0	1	1
R89	1	0	1	0	1
R81	0	1	0	1	1
R84	0	1	0	0	0
R82	1	0	0	0	0
F98	0	0	0	0	0

Figure 54. The Data Matrix for Problem 2.4.

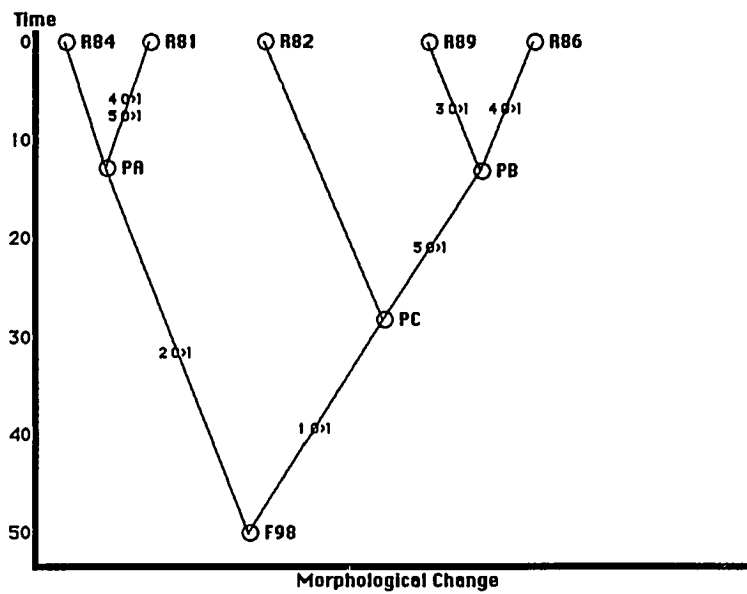


Figure 55. Solution 1 of Two Equally Parsimonious Solutions for Problem 2.4.

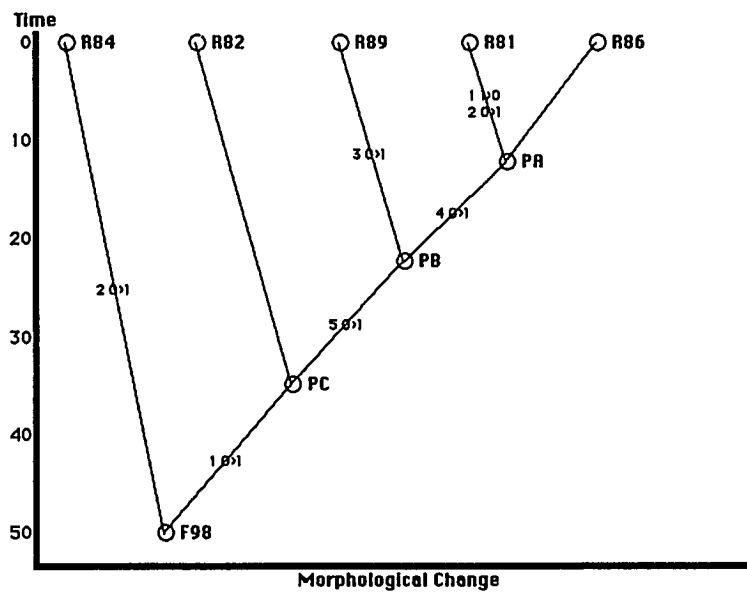


Figure 56. Solution 2 of Two Equally Parsimonious Solutions for Problem 2.4.

Problem 5.3 (Figure 61) has no evident whole ingroup synapomorphy. Characters 3 and 6 are identical and characters 8 and 10 are identical. Character 5 is exclusive from character 1 with the exception of R87. Character 5 includes nested

SPC	1	2	3	4	5
R86	1	0	0	1	1
R89	1	0	1	0	1
R81	0	1	0	1	1
R84	0	1	0	0	0
R82	1	0	0	0	0
R88	1	1	1	1	1
F98	0	0	0	0	0

Figure 57. The Data Matrix for Problem 2.4a.

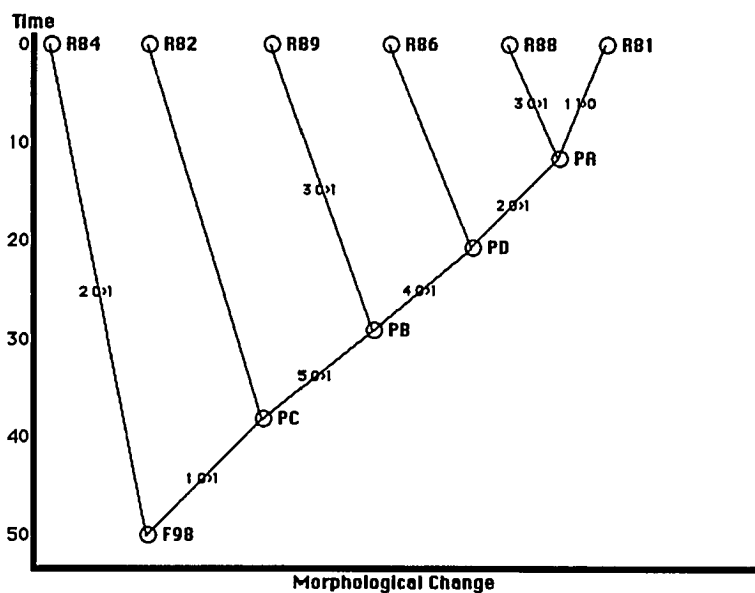


Figure 58. The Most Parsimonious Solution for Problem 2.4a.

characters including: character 7 (which splits off R89), character 2 (which splits off R88), character 4 (which splits off R87), and, with some homoplasy, characters 8 and 10 (which split off R86 and appear before the common ancestor of R82 and R80). Character 1 has characters 3 and 6 nested within and, with some homoplasy, character

9 (which appears before the common ancestor of R81 and R84). Characters 1, 8, 9 and 10 require homoplasy. Character 1 must be gained homoplasiously in R87. Characters 8 and 10 must be gained homoplasiously in R84. Character 9 must be gained homoplasiously in R80. The single most parsimonious solution Figure 62) requires 14 steps.

SPC	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
R89	0	1	0	1	1	1	1	1	0	0	0	1	1	1	1	1	0	0	0	0
R80	1	0	0	1	0	1	0	1	1	1	0	0	1	0	0	1	0	1	0	1
R81	1	1	1	0	0	0	0	1	0	0	1	1	0	0	0	0	1	1	0	0
R82	1	0	0	0	0	1	1	1	1	1	1	0	1	1	0	0	0	0	1	0
R85	0	1	1	0	1	0	0	1	1	0	0	0	0	0	1	0	1	0	1	1
F98	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Figure 59. The Data Matrix for Problem 4.3.

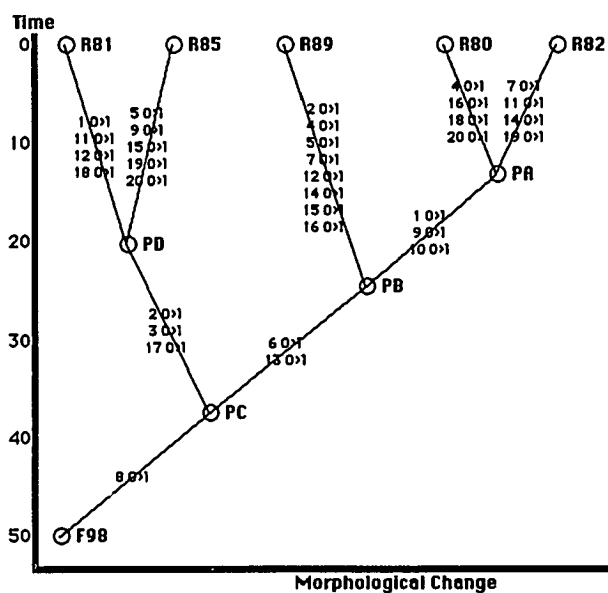


Figure 60. The Most Parsimonious Solution for Problem 4.3.

The problem set as constructed contained a number of shortcomings. All of the problems show complete resolution (i.e. result in dichotomous branching). An

axis of degree of resolution might have been another area to examine. The characters in problems 1.3, 1.4, and 1.5 were not sufficiently randomized. Some participants may have not approached these problems independently. Also, Problem 1.2, which

SPC	1	2	3	4	5	6	7	8	9	10
R81	1	0	0	0	0	0	0	0	1	0
R82	0	1	0	1	1	0	1	1	0	1
R83	1	0	1	0	0	1	0	0	0	0
R84	1	0	0	0	0	0	0	1	1	1
R85	1	0	1	0	0	1	0	0	0	0
R86	0	1	0	1	1	0	1	0	0	0
R87	1	1	0	0	1	0	1	0	0	0
R88	0	0	0	0	1	0	1	0	0	0
R89	0	0	0	0	1	0	0	0	0	0
R80	0	1	0	1	1	0	1	1	1	1
F98	0	0	0	0	0	0	0	0	0	0

Figure 61. The Data Matrix for Problem 5.3.

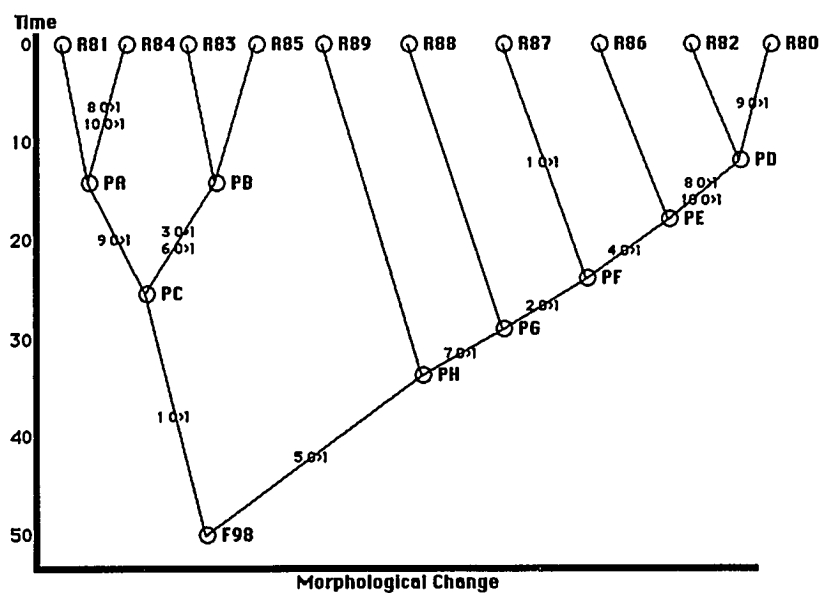


Figure 62. The Most Parsimonious Solution for Problem 5.3.

was intended to show only a character optimization, can be seen as requiring two topologies, depending on the role F98 is assumed to play with respect to the ingroup. Finally, in retrospect, it appears that complexity in phylogenetic problems is more complex than originally conceptualized.

Complexity, as originally conceptualized, was the result of increasing numbers of solutions, characters, and taxa. Each of these factors contributes to the depth or breadth of search required during problem solving. The number of solutions determines the depth of search during solution assessment. The number of characters determines the depth of search required during character assessment. The number of taxa determines the breadth of search during character assessment. In addition, however, the amount of homoplasy greatly contributes to complexity of problems irrespective of the number of solutions. One of the most commonly used measures of homoplasy is the consistency index (CI): the ratio between the number of transitions and the number of characters on a tree. This value, however, is the same across all topologies, and some topologies were found much more often than others. A better measure of the difficulty in finding solutions may be the ratio of non-homoplasious characters to potentially signal bearing characters (total number of characters minus autapomorphies and whole ingroup synapomorphies). This topic is elaborated below (page 192).

Data Gathering

Before the start of data gathering, the problem-solving environment was reviewed by two expert phylogenetic biologists for content and structure and the proposal was approved by the Western Michigan University Human Subjects Institutional Review Board (Appendix B). Data gathering began on March 9, 1995 and continued until November 22, 1995. Data analysis began as soon as the first set of

data had been collected, was conducted concurrently with data gathering until the last participant's data had been collected, and continued until March, 1996.

Participants

Recruitment efforts for this study focused on individuals with advanced degrees, who used phylogenetic analysis as a regular part of their professional occupation. Although an effort was made to find individuals across research traditions, all of the participants recruited belonged to the phylogenetic cladistic tradition. Seven of the nine participants were male and all but one participant was Caucasian. The subjects varied with respect to age and position. Three were relatively young and just embarking on their careers, three were middle-aged and well-established in their careers, two were in administrative positions past the mid-points of their careers, and one was an emeritus scholar.

Participants were initially recruited to participate in the project by posting an article (Appendix C contains the article) describing the project and the need for participants to several internet discussion groups (bionet.molbio.evolution, sci.bio.evolution and the Darwin-L mailing list), publicizing a world wide web site with the a description of the project and copies of the research software, and through targeted email to several researchers who seemed particularly appropriate. Two participants, both doctoral students at a large museum, volunteered via email as a result of this posting. The first participant suggested two people, faculty members of a nearby university, who might be good candidates. Both were contacted via email (Appendix D contains a sample message) and one agreed to participate. I then generated a list of potential participants by reviewing the relevant literature and contacted each directly via email. One did not respond. Another responded and indicated interest, but said they were too busy. Another expressed initial interest, but

subsequently failed to reply to email. I then contacted a faculty member at a university, who agreed to participate and who subsequently suggested someone as a contact at a second large museum and several other potential participants. I emailed the contact person who agreed to participate and who also suggested several more people at the same institution. Seven people at this museum were contacted by letter and 3 agreed to participate. One suggested an alternate, a postdoctoral associate at the museum, who was contacted via email and who then agreed to participate.

In the Results, each participant is referred to by a number (1-9). When referred to singly, this number will be preceded by the letter "S", for Scientist, e.g., S1. When referred to in groups, the numbers will be contained in parentheses after the term participants, e.g., participants (1, 2, 3). The descriptions below have been arranged randomly to conceal the identities of individuals.

One participant was an emeritus faculty member at a major university. This person was active in phylogenetic research in botany and had a long record of publication in phylogenetic biology.

Another participant was a faculty member of a large, prestigious natural history museum. This person was not currently active in phylogenetic research, but had published many articles in phylogenetic biology both of original research and of a more conceptual nature.

Another participant was a doctoral student studying at a major university. This person was conducting phylogenetic research in zoology toward their dissertation at a large museum of natural history. This person had published at least one article in phylogenetic biology.

Another participant was a recent graduate of a doctoral program at a major university. This person had conducted phylogenetic research for their dissertation and

had been accepted for a postdoctoral research position at a large, prestigious museum to conduct phylogenetic research in zoology.

Another participant was a faculty member at a major university. This person was active in phylogenetic research in zoology, had published many articles and collaborated on several books in phylogenetic biology, including an important methods text, and was a founding member of a phylogenetic professional society.

Two participants were zoology department chairs at a large, prestigious natural history museum. Both had long records of research and publication in phylogenetic biology.

Another participant was a postdoctoral research associate at a large, prestigious natural history museum who had conducting phylogenetic research in botany for their dissertation and was currently active in systematic, although not primarily phylogenetic, research.

Another participant was a faculty member of a large, prestigious natural history museum. This person was active in phylogenetic research in botany, had published many articles and collaborated on several books in phylogenetic biology, including an important methods text, and was a founding member of a phylogenetic professional society.

Procedures

A checklist was created to define and standardize the procedures used for each research session (See Appendix E). After setting up the equipment, participants were asked to sign a Letter of Consent (See Appendix F). I then demonstrated the program by presenting a model problem (after the first subject, I consistently selected Homoplasy 3), asked the participant what moves he or she would make to solve the problem, and then demonstrated how to perform those moves using Phylogenetic

Investigator. In these demonstrations, I avoided prescribing how the participant should solve problems or conceptualize the goal state. Participants that were uncomfortable using the computer problem-solving environment were encouraged to use paper and pencil for problem solving. After demonstrating PI, I read a set of directions (Included in Appendix F: Research Checklist) that was adapted from Larkin and Rainard (1984).

After completing the practice, a series of problems was provided and the participant solved each in turn, using either Phylogenetic Investigator or paper and pencil, and their think-aloud remarks were tape-recorded. Phylogenetic Investigator automatically time stamps and records each action taken within the problem-solving environment. Each topology generated within PI was exported as a file before other topologies were constructed. After problem solving, all notes and drawings were collected.

The problems were presented in the same order across participants. Several factors were considered in developing the order of problem presentation. It was unclear how much time would be required to complete the problems and so each series was ranked in terms of importance. Series 1, 4 and 5 were considered most important, followed by series 2 and 3. Experience with the first subject showed that it was unrealistic to attempt to use series 3. Series 4 and 5 were particularly time consuming and were reduced to single problems. S1 had proven to be capable of solving the most complex examples from those series (4.3 and 5.3), so they were selected. Series 1 was presented in a graded order based on the number of topologies in the solution. Each problem in series 2 represents qualitatively different phenomena and so the order of those problems was not considered significant, although what was considered to be the easiest problem from series 2 was placed first to introduce the revision problem type. The subjects were not informed about the rationale for how the

problems were ordered. Speculating about the order of the problems while solving problem 2.3, S7 said, "I think you do these so you that you have a hard one and then you have a little break." Other than recognizing overt features of problems (e.g. difficulty or numbers of characters and taxa), no subject made comments that implied recognition of the underlying dimensions. There was no evidence that hypotheses about the order of problem presentation was used by any subject to constrain the problem space.

There was variability among participants with respect to problems attempted and problems completed (Table 6). Participant S1 attempted and completed series 1, 2, 4, and 5. Afterwards, series 4 and 5 were reduced to single problems (4.3 and 5.3). S2 attempted and completed 1.1, 1.2, 1.3, 1.4, 2.1, 2.1a, and 2.2. S2 attempted, but did not complete 4.3, 4.1, 5.1, and 2.2a. Participants (3, 4, 5, 7) attempted and completed series 1, series 2, 4.3, and 5.3. Participants (8, 9) attempted and completed series 1, series 2, and 5.3. Participants (6, 8, 9) attempted 4.3, but did not complete asolution. Participant S6 attempted and completed series 1 and 5.3, but sufficient time was lacking to attempt series 2.

If a participant paused in their think-aloud protocol during problem solving, the experimenter would prompt the participant to continue speaking. Initially a phrase suggested by Larkin and Rainard (1984) was used: "Can you tell me what you're thinking?" This proved to elicit introspective comments. Subsequently the experimenter used the phrase, "Keep talking?" as a less intrusive way to encourage participants to continue thinking aloud without eliciting introspection.

In some cases, especially in the case of clerical activities and with respect to clerical errors, the experimenter intervened in problem solving. For cases in which participants (2, 4, 5, 7) consistently redescribed data the experimenter facilitated the redescription process to reduce clerical errors. Similarly, when a participant's spoken

Table 6
Problems Attempted (A) and Completed (C) by Participant

Participant	Series 1					Series 4					Series 5					Series 2				
	1.1	1.2	1.3	1.4	1.5	4.1	4.2	4.3	5.1	5.2	5.3	2.1	2.1a	2.2	2.2a	2.3	2.3a	2.4	2.4	
S1	AC	AC	AC	AC	AC	AC	AC	AC	AC	AC	AC	AC	AC	AC	AC	AC	AC	AC	AC	
S2	AC	AC	AC	AC	A	A	A	A	A			AC	AC	AC	AC					
S3	AC	AC	AC	AC	AC			AC			AC	AC	AC	AC	AC	AC	AC	AC	AC	
S4	AC	AC	AC	AC	AC			AC			AC	AC	AC	AC	AC	AC	AC	AC	AC	
S5	AC	AC	AC	AC	AC			AC			AC	AC	AC	AC	AC	AC	AC	AC	AC	
S6	AC	AC	AC	AC	AC			A			AC									
S7	AC	AC	AC	AC	AC			AC			AC	AC	AC	AC	AC	AC	AC	AC	AC	
S8	AC	AC	AC	AC	AC			A			AC	AC	AC	AC	AC	AC	AC	AC	AC	
S9	AC	AC	AC	AC	AC			A			AC	AC	AC	AC	AC	AC	AC	AC	AC	

Note. The problems are ordered as they were presented to participants.

remarks clearly conflicted with their actions, the experimenter brought the conflict to the attention of the participant and requested clarification.

After completion of each problem, the experimenter asked clarifying questions and requested that the solver explain the significance of various actions and provide rationales for the use of particular strategies or heuristics. In addition, at the end of the problem-solving session, the experimenter asked the participant whether the kind of thinking that they engaged in through the use of PI was a realistic characterization of their expertise in phylogenetic tree construction. Those remarks were summarized separately from the problem-solving data to evaluate the validity of Phylogenetic Investigator.

Data Analysis

Four types of research data were collected with each participant: (1) transcripts of the of the think-aloud protocol; (2) a list of actions in the problem-solving environment, time stamped and recorded by the computer; (3) any notes or drawings generated by the solver; and (4) any comments from the retrospective interview. The transcripts, actions, and notes were considered in artificial units called frames. A frame consisted of that portion of each of the three types of data associated with a unique inference in the problem-solving process. A unique inference represents a feature of either the matrix or tree that the participant referred to as the basis for taking some action. Actions initially were seen as either moving lines in the matrix or drawing some element or elements of a phylogenetic tree. In the course of analysis, actions came to include restructuring hypotheses prior to engaging in construction or reconstruction of a phylogenetic tree. The frames were then examined for evidence of the use of three major categories of problem-solving heuristics identified by Reif (1983): data redescription, solution synthesis, and solution assessment. Initially,

frames were coded according to the rational analysis. During the course of data gathering and analysis more detailed models were developed and used to interpret new data and to reinterpret the earlier data. These more detailed models facilitated identifying strategies and heuristics. Each frame was studied for evidence of: the application of content knowledge, the use of strategies and heuristics, and forms of knowledge organization.

The data from participant S2 were not included in the model. This participant approached the problems in a fashion that was fundamentally different from the other participants and his data were not directly comparable. Therefore, the results for S2 are described separately.

Development of a Model of Expert Performance

One goal of this research was to construct a model that embodied the most consistent and regular aspects of experts' strategic knowledge. This model was conceptualized as an organized network of procedures that could bring together the consistent activities used by each participant and the common activities used by all. The model was constructed by seeking to explicitly relate verbal behaviors and problem-solving activities to elements and properties of the task (van Someren, Barnard, & Sandberg, 1995). The components of such a model do not necessarily correspond directly with the psychological components of experts' mental models, but this should not be necessary for the purposes of describing desired performance.

The transcripts were collected over a period of 10 months. As each set of transcripts was transcribed, it was studied using the most recent model to evaluate the explanatory power of the model to describe the phenomena occurring within the transcripts. The set of condition action statements comprising the rational analysis

was used as the first model. Each revision extended the model to account for the regular and consistent components of problem solving that each participant used.

The model was elaborated through approximately four major iterations in order to produce a final account of expert phylogenetic tree construction. All models after the rational analysis were organized as flow diagrams. The order in which components were used was not entirely consistent, but some important relationships between components seemed best described in this form. The main features of the second model were: (a) a distinction between reasoning occurring in the matrix versus in the tree; (b) a component for building "character compatibility" (inclusion/exclusion) hypotheses involving comparison of characters and recognition of identical and "compatible" (inclusive/exclusive) characters; and (c) a distinction between tree building, character optimization, and topological optimization. The subsequent model: (a) expanded on the notion of matrix versus tree-based reasoning; (b) elaborated the structure of "character compatibility" (inclusion/exclusion) hypotheses as groups of characters; (c) developed the idea of the role of conflicting characters in the evaluation of "character compatibility" (inclusion/exclusion) hypotheses; and (d) described the model component involved in translating those characters into taxa.

The final model still makes distinctions between inferences occurring in the matrix and the tree, but makes more important distinctions based on what kinds of hypothetical entities are used to guide tree construction. "Consider inclusion/exclusion" is one of three components that deal with knowledge representations. The term "inclusion/exclusion" was used to replace "character compatibility" to improve clarity and avoid conflicts with previous uses of that terminology in the systematics literature. The term "character compatibility" has been used to describe a parallel cladistic methodology based on finding trees solely on a

criterion of "character compatibility", rather than parsimony. In addition, this form of character compatibility does not rest purely upon the recognition of inclusion/exclusion patterns: other factors can result in character incompatibility (Meacham & Estabrook, 1985). The other two components that deal with hypothetical entities describe the role of parsimony and the role of order in the matrix. Model components associated with comparisons among taxa as well as characters have been added. A section of the model now accounts for restructuring the matrix during problem solving.

The final model was used to evaluate differences across problem solvers and across problem types with respect to content knowledge, strategic knowledge, and knowledge organization. In addition, this model was also used to distinguish between general strategies consistently used by at least one participant as a predominant feature of their problem solving (which were incorporated in the model) and less generally used heuristics.

Intercoder Reliability

The final model was evaluated for intercoder reliability with the help of an associate. The associate, Dr. Robert Hafner, had participated in the development of the model and was familiar with the problems and the nature of the problem solving.

Two sets of protocols were assembled. Both sets contained transcripts in which participants had used the duplicated taxa strategy (S7), the consider order strategy (S8), and the inclusion/exclusion strategy (S1 and S4). Participants S7 and S8 were selected because they were the only participants to use their respective strategies. Two-thirds of the participants (1, 3, 4, 5, 6, 9) used the inclusion/exclusion strategy. S1 and S4 were randomly selected to provide transcripts illustrating that strategy for evaluating intercoder reliability. The first set contained the protocols from

solving problem 1.2 and the second from problem 2.1 These problems were selected because they represent nearly identical problems (5 taxon problems with 1 homoplasious character).

Using the first set of protocols, the associate and I applied the components predicted by the model until agreement was reached regarding how each component was to be applied. Subsequently, the second set of example protocols was coded by each coder, compared, and the percent correspondence calculated (73%). Afterwards, the differences were evaluated and, where possible, reconciled. Three points of difference resulted from divergent interpretations regarding instances in which the participant appeared to jump from one section of the model to another. These differences were resolved by reaching agreement about consistent application of the model. Some difficulties regarding the application of the "consider order" component of the model could not be resolved based on this sample of data. After reconciling differences, the percent correspondence was recalculated (93%).

CHAPTER IV

RESULTS

Introduction

This study used Phylogenetic Investigator to describe the content knowledge, strategic knowledge and organization of knowledge in expert phylogenetic tree construction. I have come to see phylogenetic tree construction as a body of methods for translating between two forms of data representation: the data matrix and the tree. These methods are the strategic knowledge. The form of this knowledge can be represented as potential courses of action that derive from insights about the structure of the relationships among taxa and characters. These insights could be considered content knowledge, but they have been described here in connection with the strategic knowledge they underlie. Using this distinction, content knowledge plays little role in phylogenetic tree construction as represented by the task used for this study, although some participants provided insight into how content knowledge might be brought to bear under actual circumstances of phylogenetic inference. Finally, experts used several methods to organize their knowledge during tree construction. The methods by which knowledge organization take place are described with the other components of strategic knowledge and then the implications of these methods and their associated knowledge structures are discussed separately.

This chapter deals extensively with problems, participants, taxa, and characters. Problem references always contain a decimal in them: e.g. problem 2.4. Participants are referred to in two ways. Singly, participants are referred to with the capitol letter "S" (for "scientist") followed by a numeral, e.g. S4. In groups, the term

"participants" followed by a series of numbers in parentheses is used, e.g. participants (1, 2, 3, 4). In cases of several groups of participants being described in a section, subsequent subgroups may be referred to using a term that indicates the relationships between groups of participants or between one subgroup and the whole, e.g., the others (5, 6, 7). Taxon designations always have either the letter "R" or "F" before a number between 80 and 98. Characters are single digits and, where unclear, the word "character" is used to clarify that the number used represents a character. Sets of characters (usually inclusion/exclusion groups) are enclosed in curly brackets, e.g. { 1, 2, 3 }. To represent the structure of inclusion/exclusion hypotheses, the characters are grouped using parentheses and the entire hypothesis is enclosed in curly brackets, e.g., { (1, 2, 4) (3, 5) }.

This chapter reports the results of the project in six major sections: (1) the performance and adequacy of phylogenetic investigator; (2) the strategic knowledge of experts; (3) the content knowledge of experts; (4) the knowledge organization of experts; (5) success, across participants and problems, and errors in problem solving; and (6) results pertaining to participant S2.

Phylogenetic Investigator

Phylogenetic Investigator (PI) was constructed as an environment to facilitate posing and solving phylogenetic problems and generating results useful for persuasion. In using PI for this study, I was concerned that it present the task in a form that would permit the participants to display expert performance. This concern primarily encompassed the nature of the problems, the presentation of the data, and the ability to represent solutions.

The participants fell into two groups with respect to their view on the adequacy of the problems. Some participants (1, 3, 4, 5, 6, 7) seemed comfortable

solving problems for which only coded and polarized character states were available. Others (2, 8, 9) felt that the lack of context regarding the characters and taxa in these problems made them seem unrealistic at times. Although only three participants (2, 4, 7) actually solved these kinds of problems by hand on a regular basis, all of the participants agreed that solving these kinds of problems represented an important aspect of phylogenetic thought. S8 particularly liked the revision problems, saying:

In some respects this is a more realistic problem because this is one of the problems we keep running into all the time. Adding additional taxa or adding additional characters. This is the sort of thing I keep on doing.

Participants agreed that it was important to understand how to solve these problems by hand even if software was typically used for analysis. Almost all of the participants (1, 3, 4, 5, 6, 7, 8, 9) indicated that in their research they typically used Swofford's (1991) software package Phylogenetic Analysis Using Parsimony (PAUP) either for the entire problem-solving process (1, 3, 5, 6, 8, 9) or to confirm solutions arrived at by hand (4, 7). S1, who was typical of participants who used PAUP to perform phylogenetic inference, said, "if you don't understand the stuff that I've been forced to do here, then you really don't understand how your trees are produced [using PAUP] either." S8 made similar statements, but qualified them saying:

the creating of the diagrams is a very valuable exercise, as pattern recognition, but the biology of it only when used for hypothesis testing, because the diagram itself, we were sitting here talking about something which was basically contentless. No more content than you have in the standard algorithms in higher mathematics.

Participants (2, 4, 7, 8, 9) preferred using paper and pencil to create initial representations over drawing them in PI. Two others (5 and 6) began by using PI, but as the problems became complex, also chose to use paper and pencil to create initial representations. Only S1 and S3 exclusively used PI for solving the problems.

Participants (1, 3, 5, 6, 8, 9) used PI's ability to reorder rows in the data matrix. Two participants (8, 9) who did not use the environment to create

representations, did reorder the matrix. S8 said, "The feature that's very useful in this program is the fact that you can sort the lines." Some participants (1, 4, 6, 8, 9) desired the ability to move columns as well as rows. S6, while solving problem 1.5 said, "And I can't move these characters. That's defeating one of the main ways I do things, you know, is I sort the data, and I can't sort it in the XY." In order to accommodate this participant, the data were exported into a word-processing program to facilitate moving columns of data as well.

The fact that several participants chose to use paper and pencil indicates that the representational capabilities of PI do not facilitate expert performance in phylogenetic tree construction. Participants chose to use paper and pencil for several reasons. Experts come to the problem-solving process with a suite of skills developed over many years and PI required them to learn new skills. Some participants (2, 8) were not familiar with the Macintosh and its interface and several more (4, 7, 9) felt uncomfortable using an unfamiliar computer program. S4 said:

I mean I wouldn't mind at some point in the future coming back and doing this with the computer, but I'd have to be more comfortable with the actual instructions for making the nodes and links and things and when I felt more fluent with that, then I could, like with MacClade or PAUP, where I feel more fluent with it, I can sit and talk and work with it at the same time, but so much of what I do involves teaching and even when I'm sitting down with some of my students and working on a research project, its still a learning environment, so I'd have to feel a little more comfortable with that.

Others (6, 9) felt the drawing environment constrained their ability to think visually. S6, describing the constrained representational medium, said, "The thing is you can't really add stuff using this thing. [It makes it hard to] sort of doodle." S9 explained it this way:

I can find it very difficult [to solve problems] sort of in mid-air. If I can have a piece of paper and a pencil to sort of focus my thinking, I find it a lot easier, even if I don't write anything on the paper.

Similarly, S5 said, "I find that using programs like that, I mean they're great, I'm not at all anti-computers, but for some things its just easier to draw it ... If I was trying to do a publication quality figure or something, I find computers very handy for that."

S7 offered a unique reason for her method of paper and pencil representations: it left a "paper trail" or record of the problem solving which made it easier to go back and check the solution and problem-solving process. S7 also indicated that she had previously solved problems by reorganizing the data matrix, but that she no longer did this:

At one point when I first started doing this, I used to cut out strips of paper and move them around, the way you can move lines around, so I did that at the beginning, but then I progressed, sort of, but I hate to use the word "progressed" but I just changed the way I did it. ... I would literally cut the matrix into strips back before, we used to do these on punch cards, you know, so it was really hard to do any kind of manipulation and so I would cut 'em out and move 'em around until I lined 'em up, but I ultimately went to this because it was easier to leave a paper trail. I can go back ... and keep track of what I did.

Although PI's representational abilities did not facilitate expert performance in phylogenetic tree construction, none of the experts faulted its ability to present data or to create realistic phylogenetic trees. The ability of participants to use a familiar form of representation when necessary, i.e. paper and pencil, provides confidence that the data gathered for this study are representative of the participants' typical performances. The implications of this research for future development of PI are discussed in the next chapter.

Strategic Knowledge

Strategic knowledge occurs in three hierarchical levels of organization along a dimension of scope of effect. Some phenomena are primarily local, affecting only a small part of the tree or total problem-solving process. Other phenomena were more global and dealt with effects at the level of the whole tree or problem-solving effort.

Those methods that were consistent procedures in at least one participant's approach, were used to construct a procedural model of phylogenetic tree construction. The individual components of this model comprise the middle level of the hierarchy of strategic knowledge. For example, the process for comparing characters (a component of the model) is one form of strategic knowledge. Above the level of components are what I have chosen to call "strategies" that represent more global approaches to problems and can often be represented as particular pathways through the model. Comparing characters is part of a strategy of using inclusion/exclusion analysis to define the structure of a phylogenetic tree. Below the level of the model components are what I've chosen to call "heuristics" that are more local in their scope of effect. These are different ways of approaching, implementing, using or combining components of the model. For example, the "compare characters" model component can heuristically be applied systematically or more opportunistically, by letting the eye range across the matrix.

The model and its components are described below. Heuristics are described in the context of their associated model components. Subsequently, I describe the important strategies, illustrate how they are employed, and explain why they work.

A Procedural Model of Expert Phylogenetic Tree Construction

The model (Figure 63) is organized into several spaces: the upper half deals with observations, inferences, and actions within the data matrix while the bottom half deals with observations, inferences, and actions within the tree environment. In between and on the right, are the hypothesis consideration components that mediate between these two spaces. "Consider order" deals primarily with the data matrix and

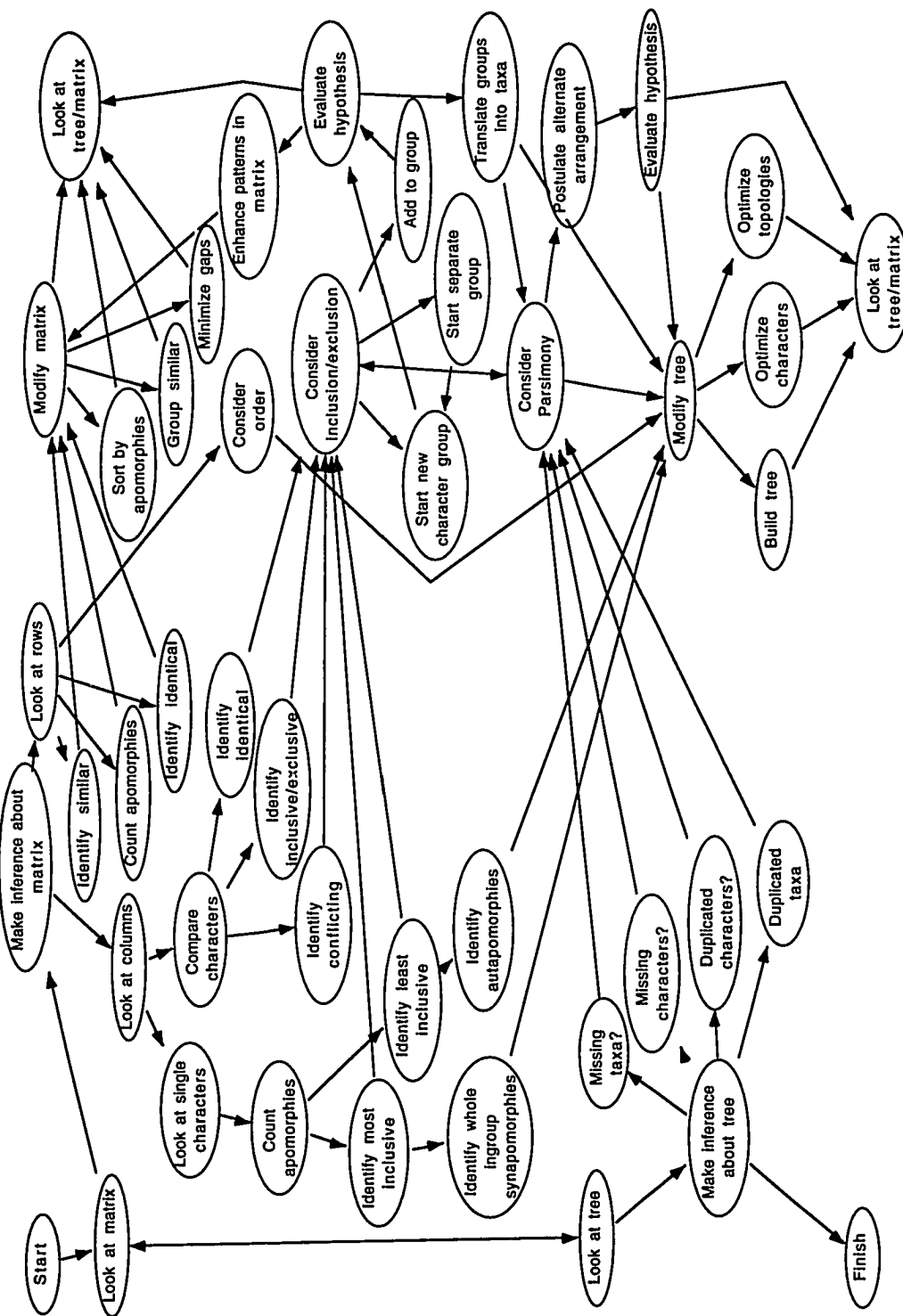


Figure 63. A Procedural Model of Expert Phylogenetic Tree Construction.

is placed near the top close to the matrix oriented components. "Consider character compatibility" uses inferences from and results in actions in both the matrix and tree environments and is placed in the middle. "Consider parsimony" takes place almost exclusively in the context of inferences in the tree environment and is placed near the bottom, close to the tree components.

As indicated in the previous chapter, the model was constructed as a flow diagram to indicate the most common pathways between components. Participants did not always generate verbalizations for every component on a pathway through the model. This might occur for two reasons: they were not processing information in a verbal form while in a component of the model or they did not use the entire pathway indicated. The pattern of verbalizations provides support for both conclusions. Pauses in verbalizations and consistent patterns of verbalization support the conclusion that participants might move through a series of components without verbalizing every one. In addition, participants often retained impressions of the tree or matrix that allowed the participant to jump directly from one region of the model to another without making a complete pass through the model from the beginning.

Components of the Model

This section provides (a) a listing of the participants that used each component of the model; (b) a brief description of the component; (c) one or two examples from transcripts that illustrate either the component being used, or an introspective statement by a participant describing its use; (d) a statement that explains how the component contributes to problem solving; and (e) descriptions of associated heuristics. The organization of the model and linkages between components represent common pathways used by participants. In fact, participants often jumped from place to place and so the model does not always work smoothly as a strict flow diagram.

Another way of representing these pathways might be a set of condition action statements as described by Larkin & Rainard (1984). I settled on the flow diagram because, although participants did not follow such patterns strictly, some patterns of flow between components did seem to be evident and the flow diagram seemed to capture the nature of these relationships most effectively.

Look at Matrix

All participants used this component. Problem solving always began by looking at the matrix. All participants went on to make inferences about relationships among characters or taxa in the matrix, except for S7 who systematically translated each character into the tree environment. The matrix contains the data for the problem and looking at the matrix is necessary for problem solving.

Make Inference About Matrix

Participants (1,3,4,5,6,8,9) used this component. Problem solving begins with inspection of the matrix. Participants (3, 5, 7, 8) tended to examine the matrix systematically: for example, S7 looked at each character individually to recode the data into individual character trees. Participants (1, 4, 6, 9) were more likely to let their eye range over the matrix looking for patterns. S4 said, "I can see patterns looking at the matrix ... I chunk it up, it's more analogous to factoring, you know look for chunks and subchunks and do those." Most participants (1, 3, 4, 5, 6, 7, 9) focused primarily upon the characters (in columns), but some participants (3, 6, 8, 9) also sought patterns among taxa (rows). Making inferences in the matrix represents describing the local phenomena of relationships among characters and taxa. The inclusion/exclusion and order of divergence tree construction strategies organize these local inferences into global hypotheses that can guide tree construction.

Look at Columns

Participants (1,3,4,5,6,8,9) used this component. Columns contain information about characters. Columns are inspected individually and compared. S9, while solving problem 4.3, recognized that she was looking sometimes at columns and sometimes at rows, and said, "I need to switch back from thinking horizontal to seeing vertical."

Looking at columns focuses attention on the evidence for structure in the tree.

Look at Single Characters

Participants (1,3,4,5,6,8,9) used this component (Figure 64). Verbalizations that illustrated inspection of individual characters were associated with systematic inspection of the matrix. For example, S5 (1.1) said, "OK, so again, its sort of the brute force thing, I just sort of go through each character." All participants made inferences based on individual characters, but did not actively examine individual characters, instead letting the eye range over the data. Identifying some single characters (whole ingroup synapomorphies and autapomorphies) can be considered immediately and therefore simplify the remaining data. In addition, the most inclusive character is often a useful starting point for character compatibility analysis.

Count Apomorphies

Only participants (3, 6, 8) systematically counted the number of apomorphies in characters. In problem 1.3, S8 said, "I'm counting the number of 0's for each character starting from number 1." The number of apomorphies indicates the number of taxa that are affected by an inference about any particular character. The relative numbers of apomorphies among characters is a factor of the importance of characters in defining the overall structure of a tree.

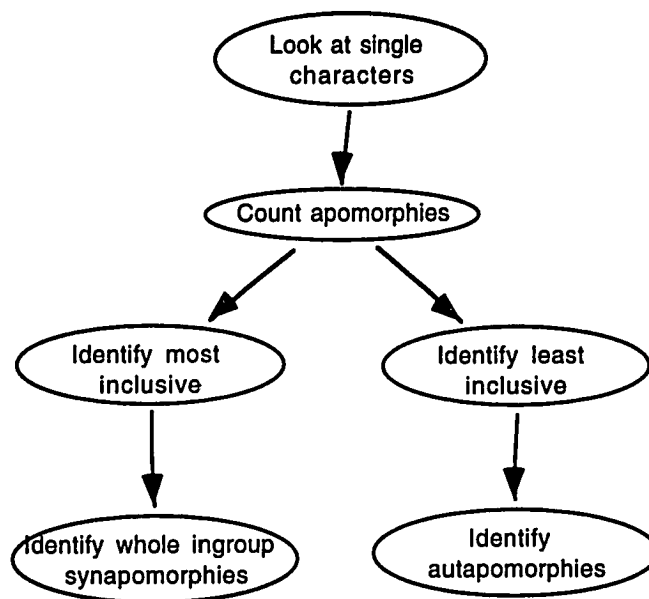


Figure 64. Model Components Associated With "Look at Single Characters".

Identify Most Inclusive

S3 used this component. S3 was the only participant who evidenced a systematic search for the most inclusive and least inclusive characters. Participants who used character compatibility (1, 3, 4, 5, 6, 9) often used most inclusive characters for beginning character compatibility groups and all participants recognized whole-ingroup synapomorphies, which is a type of most inclusive character. In this case S3 was solving problem 5.3 when he said, "The most inclusive is character 5." The most inclusive character defines the largest possible set of taxa without homoplasy and represents a natural starting point from which to make an assessment of whether the other characters support or reject this set.

Identify Whole-Ingroup Synapomorphies

Participants (1, 3, 4, 5, 6, 8, 9) used this component. Whole ingroup synapomorphies do not contribute information to determining the branching pattern within the ingroup. S8 expressed this clearly while solving problem 4.3, saying, "Character 8 is an apomorphy for the whole set and thus is trivial in separating them. It doesn't do anything for you." A whole ingroup synapomorphy can be translated directly into the drawing environment as a node linked to the ancestral node with a forward transition for the character on the link. This translation can occur immediately because no other character can conflict with this inference.

Identify Least Inclusive

Only participant (3), like with "Identify most inclusive", systematically sought this feature. For example, in problem 5.3 S3 said, "then for the least inclusive, character 3 puts [R] 83 and [R] 85, character 6 puts [R] 83 and [R] 85 together, character 8 puts [R] 82 and [R] 84 together." Although no other participants made verbalizations associated with seeking or finding least inclusive characters, all the participants recognized autapomorphies, which are a subset of least inclusive characters.

Identify Autapomorphies

Participants (1, 3, 4, 5, 6, 7, 8, 9) used this component. All the participants recognized autapomorphies as characters which were useless for determining the branching patterns within the ingroup. Solving problem 4.2, S1 said, "Oh! Here's a useless character — character 4, we can ignore that for the rest of the discussion."

That's nice." Autapomorphies, being uniquely derived characters, cannot conflict with other characters, and are always represented as a transition immediately prior to the taxon which possesses it.

Compare Characters

Participants (1,3,4,5,6,9) used this component (Figure 65). Participants (3, 5, 6) appeared to actively hold single characters in mind while comparing them to all others. For participants (1, 4, 9) this process seemed to occur as a very general pattern recognition activity where the impression of a sequence of 1's and 0's would be held in mind as the eye passed over the rest of the data. The particular features that were most often recognized are described below. S9, solving problem 1.5, made this introspective comment: "I think the first thing I'm doing when I look at these sets of characters is actually looking for any groups of characters that are congruent. And/or identical."

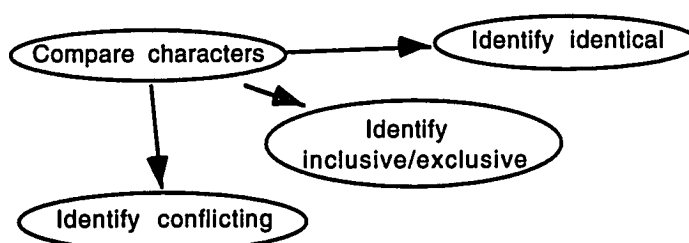


Figure 65. Model Components Associated With "Compare Characters".

When there were a large number of characters (especially problem 4.3), two participants (4, 5) created a list of the taxa claimed by each character and as each new character was added, they compared it with the previously considered characters for identity and compatibility. This heuristic greatly facilitated the construction of a

character compatibility hypothesis. Only one other participant (S1) found a correct solution to this problem.

Identify Identical

Participants (1,3,4,5,6,9) used this component. Identical characters have the same distribution of 1's and 0's. Solving problem 1.5, S6 said, "I'm looking for patterns of identity in them. Anything that makes the problem to be of less generality. Either characters or taxa that are near duplicates of one another." Identical characters were almost universally sought by the participants. Recognizing identical characters not only identifies strong support for a single node, but it also simplifies the number of patterns that need to be accounted for during hypothesis construction. Identical characters were used by several participants (1, 4, 5, 6, 9) as the basis for beginning an inclusion/exclusion hypothesis. (S7 also identified identical characters, but did so by comparing the set of taxa defined by the characters with already drawn character trees.)

Identify Inclusive/Exclusive

Participants (1, 3, 4, 5, 6, 9) used this component. Characters are inclusive when the distribution of 1's for one character is a subset of another character and exclusive when they describe mutually exclusive groups. In this example from problem 1.2, S1 observes, "the characters 1, 2, 4, and 5 are compatible with each other." Participants often used the term "compatible" to refer to whether or not two characters were inclusive or exclusive with respect to one another. For the purposes of describing this relationship, the term "inclusive/exclusive" is used because the term "compatible" has a prior history of use in the systematics literature that is different from the use here. Characters that were exclusive, but that between them did not

account for the entire ingroup were often identified as inclusive/exclusive, but in some situations, were overlooked.

Identify Conflicting

Participants (1,3,4,5,6,9) used this component. Characters conflict when they define groups of taxa that incompletely intersect. Character conflict was evaluated at two levels: the first involved conflict between individual characters and the second was between characters and an inclusion/exclusion hypothesis. In this example S3 describes conflict between characters:

There's conflict between character 1 and character 4 which I would first guess would probably make 2 parsimonious reconstructions, 2 equally parsimonious reconstructions dealing with those two. Nothing conflicts with character 2.

Characters that conflict with the current inclusion/exclusion hypothesis must be explained either through homoplasy or as part of an alternate hypothesis. Evaluating the extent to which other characters conflict with an inclusion/exclusion hypothesis is the basis upon which they are evaluated.

Look at Rows

Participants (3,6,8,9) used this component (Figure 66). Participants (3, 9) used a quick examination of rows as a heuristic to reorganize the matrix to facilitate pattern recognition of characters. Only S8 and, to a lesser extent S6, used this as a primary means of making inferences about phylogenetic relationships.

Identify Identical

Participants (6,8,9) used this component. For some (6, 8), this represented a fundamental flaw in the data: if you can't tell taxa apart then they aren't really separate

taxa. Typically, this information was used to organize the matrix. S8, while solving problem 1.1, said:

the next thing I'm looking for is there are basically two identical here, there are two identicals, R81 and R86 are identical, there is a 10111, 10111, so they are identical taxa, so you can cut those out.

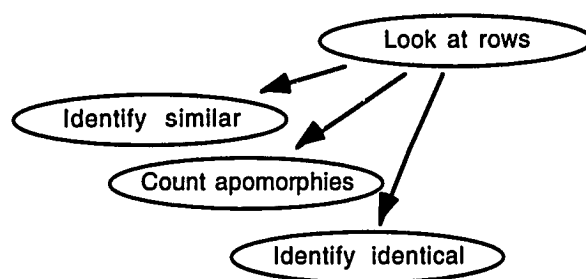


Figure 66. Model Components Associated With "Look at Rows".

The recognition of identical taxa has direct implications for the tree. No other evidence can contradict the fact of identical taxa and so this inference can be acted upon immediately. Identical taxa will always be sister taxa sharing a recent common ancestor. When constructing a phylogenetic tree, S8 often didn't make separate branches for identical taxa. They would be placed together at the end of a single line to indicate that they could not be distinguished.

Identify Similar

Participants (3,6,8,9) used this component. Those participants who actively reorganized the matrix often placed similar taxa together as well as identical taxa. While solving problem 1.5, S8 said, "These two, [R] 83 and [R] 88 in this case, are very similar and are separated only by an apomorphy in character 1. Similar taxa will usually appear as sister taxa in the most parsimonious tree.

Count Apomorphies

Only participants (6,8) used this component. The number of apomorphies per taxon was typically used to reorganize the order of taxa in the data matrix: taxa having many apomorphies were placed at the top and taxa with few at the bottom. When homoplasy is low, the number of apomorphies will be strongly correlated with the order the taxa will have diverged from the outgroup. In this example from problem 5.3, S8 is counting apomorphies:

R81 has only 2 apomorphies. 2 has ... 6, I'm assuming there are some in between. [R] 83 has ... 3 [R] 84 has ... 4. [R] 85 has ... 3. [R] 86 has ... 4 [R] 87 has ... 4. [R] 88 has 2. [R] 89 has 1. OK. [R] 80 has ... 7, what am I saying? [R] 80 has 7.

Modify Matrix

Participants (1,3,5,6,8,9) used this component (Figure 67). In the version of PI used for this research, the only movements that could be made in the matrix were repositioning rows. S6, while solving 1.5, said "One of the main ways I do things, you know, is I sort the data."

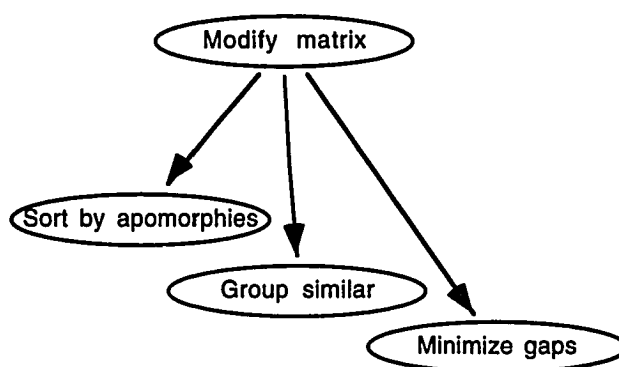


Figure 67. Model Components Associated With "Modify Matrix".

Participants (1, 6, 8, 9) expressed an interest in being able to reorder characters as well. For example, S6 said, "Can I sort rows and columns in the matrix here or not?" Reordering rows facilitates using the matrix to represent hypotheses of relationship and evaluating differences between taxa.

Sort by Apomorphies

Participants (3, 6, 8) used this component. This component has two levels of consideration. S8 counted the number of apomorphies systematically and tended to organize the matrix with the taxa with the most apomorphies at the top. For example in problem 1.3, S8 said, "[R] 81 must be the pinnacle." Those with the fewest were placed at the bottom. For example, S3 said while solving problem 2.1, "I consider [R] 89 to be the most primitive." Sometimes participants would move taxa to the top or bottom because they "looked" apomorphic or plesiomorphic. When homoplasy is low, the number of apomorphies will be strongly correlated with the order the taxa will have diverged from the outgroup.

Group Similar

Participants (3,6,8,9) used this component. Taxa which had identical or just similar distributions were often placed together. S9 said while solving problem 5.3, "[R] 80 move that to between [R] 89 and [R] 82, yes, I thought so, its very similar to, [R] 80 and [R] 82 are very similar." Although participants made clear distinctions about finding identical or similar taxa, they tended to act the same in both cases with respect to organizing the matrix. Identical taxa, and usually similar taxa, are sister taxa and grouping them together serves to organize or "chunk" the data.

Minimize Gaps

Participants (3,8) used this component. This component represents efforts to reorganize the matrix so that as many characters in the matrix as possible have all of their 1's together. S3 said, while solving problem 5.3, "I'm just using a, let's-minimize-the-gaps-in-the-matrix sort of approach." The result of minimizing gaps is that taxa in the matrix are organized to reflect the largest inclusion/exclusion groups.

Consider Order

Participants (6,8) used this component. S8, and in one example S6, used the order in which taxa appeared in a restructured data matrix as the primary means by which a phylogenetic tree would be constructed. While solving problem 1.1, S6 said, "I like to this solve this stuff often by not even thinking about steps, just thinking about who goes next to who." After reorganizing the taxa, those lowest in the matrix would be the first to branch away from the ancestor and those at the top would be placed at the upper right-hand part of the tree and characters would be mapped onto this tree as more of a justification rather than as the primary motivating factor. When homoplasy is low, the correspondence between the number of apomorphies of a taxon and its relative position on the tree will be very high and this strategy will work well.

Consider Inclusion/Exclusion

Participants (1,3,4,5,6,9) used this component (Figure 68). This component describes how participants constructed and used hypotheses about how characters fit together. These hypotheses involve seeking sets of characters which unambiguously support discrete features (postulated nodes and links) of a phylogenetic tree. Grouping

characters organizes, or "chunks," the data and facilitates assessment of the number of characters that support a hypothesis and those which contradict it.

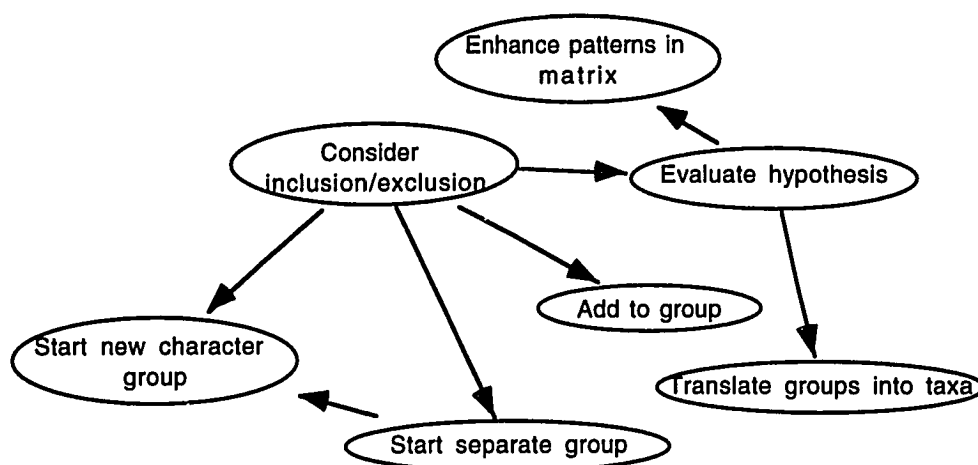


Figure 68. Model Components Associated With "Consider Inclusion/Exclusion".

Inclusion/exclusion hypotheses appeared to be composed of two or three groups of characters. A participant would start by finding a most inclusive character or pair of identical characters and then place nested characters within that group. If the participant found an exclusive character, it would be used to start a separate group. Conflicting characters would be placed in a third group. Once all the characters had been considered, a hypothesis was evaluated by considering how many characters it could explain without conflict and comparing this to the number of conflicting characters. In general, the hypothesis capable of explaining the largest number of characters without homoplasy was selected.

In some cases, hypotheses of inclusion/exclusion were global and represented complete surveys of the data set. In this example from problem 2.4, S5 identified the two basic trees to be constructed:

characters 1 and 2 agree with each other, character 3 agrees with anything basically, character 4 doesn't agree with 1 or 2 character 5 agrees with 4 but not with 1 or 2, so here its hard to pick a sort of a set of characters that think you ought to just go with because there are actually sort of 2 sets and its basically 1 and 2 vs. 4 and 5 and so what I might actually do is draw trees based on those two sets initially and see what they say about each other.

Often, however, inclusion/exclusion hypotheses were relatively incomplete and once a local pattern was found, it was translated into taxa and used to begin tree construction. For subsequent construction the participant would sometimes return to further evaluate inclusion/exclusion, but often would consider parsimony to evaluate where to place missing taxa and characters. S1, while solving 2.1 said:

Characters 1 and 3 they're compatible with each other, but they got outvoted. ... Characters 1 and 3 ... form a clique which is incompatible with the biggest clique, yeah, can't be most parsimonious, so we have to fit those onto this tree. Character 3 postulates [R] 88 plus [R] 85 plus [R] 86, oh that puts [R] 85 and [R] 86 together and character 1 puts [R] 85 and [R] 86 together, oh I see, character 1 is compatible with everything else

Start New Character Group

Participants (1,3,4,5,6,9) used this component. A new group started when a participant noticed that characters were either identical or nested. S1, while solving problem 2.3, said, "I see character 2 and 4 are identical this time and they happen to be compatible with character 3. And with character 5." Characters 2 and 4 are being used to start a new character group, and subsequently 3 and 5 are being added to it. Grouping characters organizes the data and allows a general assessment of the number of characters that support a hypothesis and those which are against it.

Add to Group

Participants (1,3,4,5,6,9) used this component. Once a group has been started, more characters can be placed within the group if they are identical to or nested within another character in the group. S1, while solving problem 2.3, said, "I see

character 2 and 4 are identical this time and they happen to be compatible with character 3. And with character 5." In this passage, characters 3 and 5 are added to a group that was started based on characters 2 and 4. Grouping characters organizes the data and allows a general assessment of the number of characters that support a hypothesis and those which are against it.

Start Separate Group

Participants (1,3,4,5,6,9) used this component. There are three useful distinctions to make among inclusion/exclusion groups: (1) characters that are identical to or nested within characters can be placed in the same group, (2) characters that are exclusive from a group should start a new group, and (3) characters that conflict with any other group should be set aside in a third group. S4, while solving problem 1.3 said, "[Character] 2 links up the two taxa that are outside of what's with 3 and 5." Starting a separate group, provides a separate "container" that can hold any further characters which might be nested within a character that is exclusive from other characters.

Evaluate Hypothesis (Inclusion/Exclusion)

Participants (1,3,4,5,6,9) used this component. Evaluation of inclusion/exclusion hypotheses are usually stated in numbers of characters. S1, while solving problem 5.3 said, "Alright, so now I've managed to explain 5 characters without homoplasy." These statements can simply report the number of characters that support or conflict with the hypothesis, or compare the number of characters supporting the hypothesis either with the number that conflicts or with the number of characters that support some other arrangement. This component determines whether

there is enough evidence to start drawing a tree. S5, while solving problem 2.4 said:

So here it's hard to pick ... a set of characters that think you ought to just go with because there are actually sort of two sets and its basically 1 and 2 vs. 4 and 5 and so what I might actually do is draw trees based on those two sets initially and see what they say about each other.

In addition to comparing the number of inclusive/exclusive and conflicting characters, participants (1, 3, 4, 5) would perform an abbreviated inclusion/exclusion analysis of the conflicting characters. A consistent pattern among the homoplasious characters indicates a high probability of other most parsimonious solutions.

Similarly, the lack of such a pattern is evidence that supports the current solution as the best solution. S1 (5.3) said, "none of these homoplasies appears to help each other." Furthermore, when conflicting characters are mutually incompatible, they add support to the hypothesis being considered. S1 described this phenomenon by saying:

characters that contradict the characters that contradict that characters you want to believe in are support for the characters you had to start, so if I see a bunch of mutually contradictory characters, that suggests I should start by ignoring those characters and look for those cliques [inclusion/exclusion hypotheses] that do support each other

Enhance Patterns in Matrix

Participants (1,5) used this component. These participants would reorganize the matrix to make it match a currently existing tree or inclusion/exclusion hypothesis, usually in the context of trying to consider other most parsimonious arrangements of taxa. While solving problem 4.3, S1 said, "To make this easier, I should put my clades together on the matrix, so I will do that, I'll move that, I should make the matrix match the structure of my trees as closely as I can." This practice allows making and comparing parallel inferences in the matrix and the tree.

Translate Groups Into Taxa

Participants (1,3,4,5,6,9) used this component. Once an inclusion/exclusion hypothesis has been evaluated positively, participants would translate the hypothesis of characters into sets of taxa. Each set of taxa would be based on a node or postulated common ancestor and supported by a character or characters. Inferences about characters are acted upon by arranging the taxa according to the sets defined by the inclusion/exclusion groups. S1 solving 1.3 said:

so I will start by constructing the parts of the topology that are unambiguous and we see immediately that R83 goes with R81 due to character 2; R87, [R] 84, and [R] 85 go together because of characters 3 and 5; and then we have only to pick one resolution with relationship to R87, [R] 84, and [R] 85.

The translation process usually resulted in having the taxa listed across the top of the drawing field (either in PI or on paper). A typical arrangement would be with the least apomorphic on the left to the most apomorphic on the right. S7 suggested handedness as a possible factor here:

I once had a student ask me if all cladograms had to be drawn off to the right, this was years ago, I said no, if you look at the books that [S4] and I ... did together, you'll see all his went to the left and all mine went to the right ... because he's left handed and I'm right handed. And he said, Oh, it makes sense, because if you drew a cladogram like this and then tried to write on it, you'd get the ink all over the place.

Consider Parsimony

Participants (1,3,4,5,6,7,8,9) used this component (Figure 69). This component acts as a filter where participants evaluate how the parsimony of a tree will be affected by placing a character or taxon on a tree. In simple situations, characters and taxa can be placed unproblematically, which supports the direct pathway to "draw tree". In other cases, the effects on other characters and taxa needs

to be considered before modifying the tree. In these cases, competing possibilities and their implications often must be considered and compared to decide which to select.

In this example S1 considers parsimony, postulates an alternate arrangement, and then evaluates the arrangement:

Alright, six characters have no homoplasy -- that we like . Four characters have only one extra step, that's not bad. I can actually save two steps by putting R84 over here, but it would cost more steps than that... considerably more steps than that.

S3 was capable of evaluating inclusion/exclusion hypotheses for parsimony.

S3 said this while solving problem 2.4:

If we use 1 and 2 as the starting point, that means that 4 has evolved twice and 5 has evolved twice if we use 4 and 5 as a starting point, that means that 1 has evolved 1, 2, 3, and that 2 has evolved twice. So I think that means that the 1, 2 is more consistent.

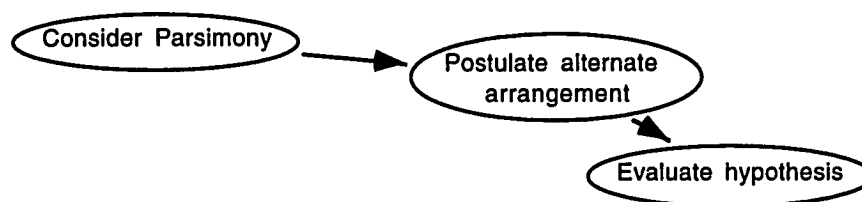


Figure 69. Model Components Associated With "Consider Parsimony".

Postulate Alternate Arrangement

Participants (1, 3, 4, 5, 6, 7, 9) used this component. This component works by running through the permutations of other ways a taxon or character could be placed on the tree and what effects this arrangement would have on the rest of the tree. Homoplasious characters often have multiple optimizations and imply multiple arrangements of taxa. Some or all of these need to be considered for parsimony.

While solving problem 2.4, S5 said:

characters 2,4, and 3 all fit this tree perfectly and characters 5 and 1 don't, so what I want to see is what groupings those suggest and whether they could improve anything, whether you could rearrange it to improve the fit of those characters.

Postulating alternate arrangements work for relatively simple rearrangements of the taxa or characters, but large-scale changes require a lot of working memory and often result in failure when the participant becomes distracted or confused.

Evaluate Hypothesis (Parsimony)

Participants (1,3,4,5,6,7,9) used this component. Once an alternate arrangement has been fully developed, it is evaluated by comparing how many steps would be saved by using the new arrangement but at what cost to other characters. S5, while evaluating an alternate arrangement of the taxa said, (2.4) "if you put it down at the base it'll require three steps instead of two." This component determines which potential arrangements of characters and taxa are most parsimonious.

Look at Tree

All participants used this component. Looking at the tree is necessary to evaluate what is missing during the construction of the tree and what could be improved on the tree during optimization.

Make Inference About Tree

All participants used this component (Figure 70). Participants needed to examine the tree for completeness during construction and adequacy during assessment. Inferences about the tree result in parsimony considerations.

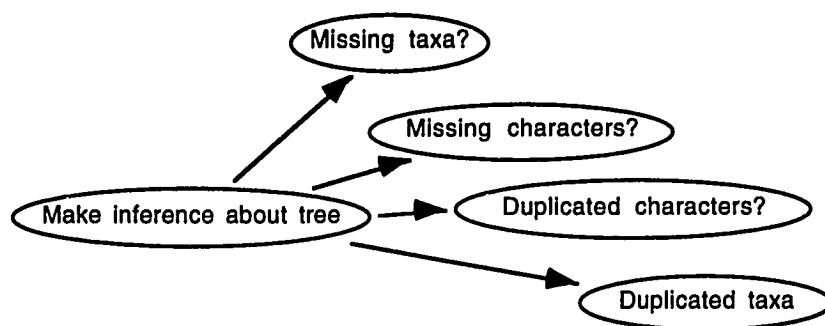


Figure 70. Model Components Associated With "Make Inference About Tree".

Missing Taxa?

Participants (1,3,4,5,6,7,8) used this component. During tree construction, participants were concerned with assuring that all of the taxa were placed on the tree. S5 during the initial construction of a solution said, (5.3) "Is that all the taxa? No, so there's two taxa still unplaced." Missing taxa are added by considering where they could most parsimoniously be added to the current arrangement.

Missing Characters?

Participants (1,3,4,5,6,7,8,9) used this component. Participants that used inclusion/exclusion (1, 3, 4, 5, 6, 9) constructed the tree based on the characters in the inclusion/exclusion hypothesis, distributing each compatible character as the structures they supported were added to the diagram. Afterwards, the homoplasious characters would need to be recognized as missing from the diagram and distributed. S1, after linking all of the taxa into a tree structure in problem 4.3, said, "And at this point, I have no idea what other characters ought to be there or ought not to be there so I guess I'll just have to go through them all and see what they look like." S1, after

placing all of the non-homoplasious characters in problem 5.3, said "OK, we're still missing character 1, 8, 9, 10."

Duplicated Taxa?

Participant (7) used this component. This component reflects the primary means by which S7 solved phylogenetic problems. S7 translated the matrix into small trees and looked for duplicated taxa and groups of taxa that could be unproblematically combined. S7, recognizing that two combination trees both claimed the same taxon, said in problem 2.1:

So if you combine these, you have a conflicting statement with [R] 88. So at this point, you've got three characters that put [R] 88 over here, one character that puts [R] 88 over there,. so I would remove [R] 88 from here and have a parallel 3 over there.

Resolving character duplication among the non-conflicting characters results in a hypothesis that is consistent with an inclusion/exclusion hypothesis (thought not necessarily the largest one). Once all of the unproblematic duplications had been considered, S7 used parsimony to resolve the situations where characters conflicted.

Duplicated Characters?

Participants (1,3,4,5,6,7,9) used this component. Characters which appear more than once on the tree are the homoplasious characters. In solving problem 4.3, S3 used one of the largest inclusion/exclusion groups (See the Inclusion/Exclusion strategy below) to construct a tree and then sought to determine whether this tree was the most parsimonious tree. S3 said:

So the only other thing is to look at consistent patterns of the same numbers showing up ... looking at the tips leading to, the various branching leading to tips, seeing if there are any characters that show up together, as being possible ones to group down below, and I don't see any, frequency of more than two apparent together, so my guess would be that this would be the most parsimonious.

He used the quantity "more than two" here because two characters, 7 and 14, support the division he selected among the three problematic taxa. There are three duplicated characters {1, 9, 10} in this example, but he was unable to recognize them and accepted a less parsimonious solution. Participant S4 used a visual metaphor to describe the attainment of the same goal. At this point in problem 1.5, there were two characters distributed homoplasiously:

if you take a situation like this and you push all these putative homoplasies way down the tree, that's the easiest way to tell if membership in these is going to collapse and you're going to have to consider moving things around, if you push them way, way down, then you're going to remove as much support above as possible and if you've made a mistake in that initial allocation, you'll see it, um much more quickly

A heuristic for this component is to consider the characters which required the most steps on the current tree first. While solving problem 4.3, S1 said:

I should now look at characters that have three steps and see if any radical rearrangement will save them something. ... so, I think I have no characters requiring three steps to explain here, and I have a number of characters requiring two steps to explain, but any other topology would probably do violence to the parsimony.

Fixing the characters which add the most length to the current tree provides the greatest likelihood of resulting in a new topology. A character that saves two steps, can break two other characters and result in an equally parsimonious hypothesis.

Modify Tree

I conceptualized tree construction as a series of modifications to the drawing environment (Figure 71). Initial steps involve the construction of the tree: Taxa were arranged across the drawing field, nodes were postulated supported by characters, everything was linked together. Subsequent steps involve optimization of the characters and topology. Both tree construction and optimization were approached in a very mechanical fashion. The thinking that guides tree construction and

optimization occurs in the components where hypotheses are considered: Consider order, Consider inclusion/exclusion, and Consider parsimony.

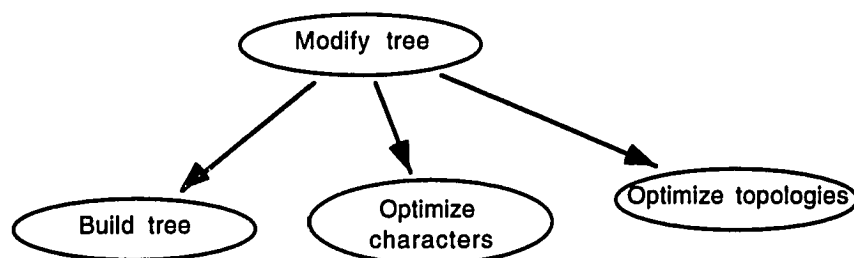


Figure 71. Model Components Associated With "Modify Tree".

Build Tree

Participants (1,4,5,6,7,8,9) used this component. The build tree component guides the initial addition of tree components, including nodes, links, and the set of transitions associated with non-homoplasious characters. S1, using Phylogenetic Investigator, said, "We start off by building a tree." S4, while drawing with paper and pencil said, "Let me take a shot at drawing this thing out." The "Build tree" component is used to tie all of the taxa together into a system of relationship usually based on a set of unambiguous characters, typically a translated inclusion/exclusion hypothesis.

Earlier versions of the model included separate components for adding links, nodes and transitions, but very little thinking seemed to take place during the construction of the tree. Activity here reflected a primarily mechanical reconstruction of hypotheses generated in other parts of the model.

There were a number of interesting heuristics associated with forms of representation, primarily associated with drawing conducted outside of PI. S7 used

different symbols to illustrate whether a transition was a synapomorphy (a single stroke), a parallel gain (parallel strokes), or a reversal (an "X"). This facilitated distinguishing between homoplasious and non-homoplasious characters. To illustrate (usually homoplasiously) shared characters, S4 would occasionally draw a single stroke across two branches. This form of representation was used as a form of shorthand which was expanded in final solutions. Finally, S7 indicated that she sometimes used dashed lines to indicate two places where a taxon might be inserted into a tree with equal parsimony. In this way a single diagram could illustrate an essential difference between two equally supported competing hypotheses.

Optimize Characters

Participants (1,3,4,5,6,7,9) used this component. Participants optimize characters by removing transitions from one place and adding others and by drawing subtrees. Here is an example of S1 solving 1.3 using Phylogenetic Investigator, "I'll take these multiple occurrences of character 4, destroy it, and take PD to PC and give it character 4 so character 4 is now parsimoniously explained and lets remove character 1 from that node."

Participants were not particularly concerned about systematically describing character optimizations. S9 echoed the general sentiment by saying, "I daresay I could manipulate a few little options with this homoplasy, but I don't think it would be worth it." Finding character optimizations was seen as something fairly simple and mechanical. Perhaps the consideration of character optimizations is not considered as important during tree construction as during the interpretation of finished trees.

Optimize Topologies

Participants (1,3,4,5,6,7,9) used this component. Participants optimized topologies by restructuring existing diagrams (using "Reassign Links") or through making new drawings that involved copying some aspects and then drawing new parts *de novo*. Solving problem 1.5, S1 said, "so I have to fix the topology to make R84 a sister group of... Sorry, change the topology to make everything except R86 a clade." In this case, S1 postulated an alternate hypothesis, considered it for parsimony and then considered the mechanical moves necessary to put it into place using Phylogenetic Investigator. Solving problem 1.3, S4 described the pattern of topologies while sketching them by hand:

... within that you have two alternative topologies, with the four character optimizations, but that devolves down into this, this, let's see, [R] 87 and [R] 85, character 1 and, with homoplasy for character 4, or here [R] 85 and [R] 84 together with character 4 and [R] 87 over here, with [R] 87 with 1 and [R] 85 with 1, so you have these two alternative topologies for that resolution and then each of these alternative topologies also has the ... the reversal option

An Example Problem

An example problem is described in detail below to illustrate the range of data that was available for analysis and show how the model was applied to the data . Table 7 contains a key for interpreting the program operation codes, their meaning, and the associated data. These codes are used in the example problem.

The example contains the complete range of data gathered with a participant using Phylogenetic Investigator: the program operations; the transcript of the think-aloud protocol; images of the matrix and drawing environment; and analysis, including descriptions of actions, references to components of the model, and interpretation of strategies and heuristics. Table 8 contains the example problem.

Table 7

A Key to Program Operations Used in Example Problem

Code	Operation	Information
OPR	Open Project	Date
OPW	Open Window	Window name
ADN	Add node	Postulated axon Designation
MON	Move node	New Location
ADL	Add link	Endpoints of new link
ACT	Add/change transition	Endpoints of selected link
MSL	Move matrix row	Old location, new location
RAL	Reassign link	Old endpoints, new endpoints
REL	Remove link	Endpoints
REN	Remove node	Postulated taxon designation

Program operations are in the Monaco typeface. Transcripts of the think-aloud protocol and comments about the participant's problem solving are in plain type. Components from the model are in boldface. Initial comments in parenthesis describe what physical effect the action had on the diagram. Comments at the end of sections interpret the rationales for actions and describe associated heuristics. This example will also serve to illustrate the most commonly applied strategy, inclusion/exclusion, which will be described in more detail below.

Table 8

Problem 5.3 as Solved by S1 Interpreted Using the Model of Expert Phylogenetic Tree Construction

Operations	Transcript of Think-aloud Protocol	Analysis and Images																																																																																																																																																																																																																																							
3:18:26 PM OPR 5/7/95																																																																																																																																																																																																																																									
3:18:26 PM OPW wd 1	<p>In this problem, S1 initially begins down an incorrect path toward the solution, but then corrects his analysis as evidence begins to mount that his initial inferences were wrong. When the participant opens the problem, he sees the data matrix and the nodes arranged in a drawing environment.</p> <p>Time 0 ○ RB1 ○ RB2 ○ RB3 ○ RB4 ○ RB5 ○ RB6 ○ RB7 ○ RB8 ○ RB9 ○ RB8 10 20 30 40 50 ○ F98</p>	<p>Tree Length=0 Problem: S1 Example 5.3</p> <table><tr><th>SPC</th><th>1</th><th>2</th><th>3</th><th>4</th><th>5</th><th>6</th><th>7</th><th>8</th><th>9</th><th>10</th><th>11</th><th>12</th><th>13</th><th>14</th><th>15</th><th>16</th><th>17</th><th>18</th><th>19</th><th>20</th></tr><tr><td>RB1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr><tr><td>RB2</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr><tr><td>RB3</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr><tr><td>RB4</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr><tr><td>RB5</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr><tr><td>RB6</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr><tr><td>RB7</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr><tr><td>RB8</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr><tr><td>RB9</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr><tr><td>F98</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr></table> <p>Morphological Change</p>	SPC	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	RB1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	RB2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	RB3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	RB4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	RB5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	RB6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	RB7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	RB8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	RB9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	F98	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SPC	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20																																																																																																																																																																																																																					
RB1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0																																																																																																																																																																																																																					
RB2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0																																																																																																																																																																																																																					
RB3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0																																																																																																																																																																																																																					
RB4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0																																																																																																																																																																																																																					
RB5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0																																																																																																																																																																																																																					
RB6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0																																																																																																																																																																																																																					
RB7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0																																																																																																																																																																																																																					
RB8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0																																																																																																																																																																																																																					
RB9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0																																																																																																																																																																																																																					
F98	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0																																																																																																																																																																																																																					

Table 8—Continued

Operations	Transcript of Think-aloud Protocol	Analysis and Images
3:18:26 PM wd 6	OPW Once again, let's try maximal clades Look for identical characters I see no identical characters wait a minute, I do.	Look at matrix Make inference about matrix Look at column Compare columns
	Characters 8 and 10 are identical. Characters 3 and 6 are identical. and characters 8 and 10 and characters 3 and 6 are compatible, so this is an excellent place to start.	Identify identical {8,10} Identify identical {3,6} Consider inclusion/exclusion { (8,10) (3,6) } Evaluate hypothesis
3:19:58 PM PA	ADN Make me a couple clades, nodes.	Rather than constructing a well-articulated character compatibility hypothesis, S1 tends to make inferences and translate them individually into the drawing environment. (Node PA is added). Build tree
3:20:01 PM PB	ADN PB	S1 starts building the tree by adding nodes, and then goes back and translates the inclusion/exclusion groups into taxa. (Node PB is added).

Table 8—Continued

Operations	Transcript of Think-aloud Protocol	Analysis and Images
3:20:10 PM MON R83 150,30		(The taxa are moved into the appropriate positions)
	Character 3 supports 83 plus 85 and character 8 and 10 support Hmm. So 83 goes with R85,	Translate groups into taxa
3:20:15 PM MON R81 525,30		Build tree
3:20:30 PM MON R82 570,30	let's put that over in the corner there.	(Move R81 and R82 out of the way)
3:20:32 PM MON R83 61,30		(R83 and R85 are brought together at the left-hand side of the screen.)
3:20:35 PM MON R85 115,30		(PB is moved out of the way)
3:20:37 PM MON PB 259,71		(PA is placed below R83 and R85)
3:20:39 PM MON PA 89,86	and we'll give that to PA R82, 84, and 80 go together. There's a node there with that.	(R82 is brought next to R84)
3:20:50 PM MON R82 253,30		(R86 is moved out of the way to the right)
3:20:52 PM MON R86 585,30		(R80 is brought next to R84 and R82)
3:20:53 PM MON R80 312,30		(R87 is moved out of the way to the right)
3:20:55 PM MON R87 476,30		Subsequent moves adjust positions to reduce overlap
3:20:57 PM MON R80 316,30		

Table 8—Continued

Operations	Transcript of Think-aloud Protocol	Analysis and Images
3:20:59 PM MON PB 254,96 3:21:05 PM MON R83 61,30 3:21:06 PM MON PA 96,83		<p>(A link is formed between PA and R83)</p> <p>(A link is formed between PA and R85)</p>
3:21:07 PM ADL PA R83 3:21:08 PM MON PA 103,82 3:21:09 PM ADL PA R85	Now I'll make some links.	

Table 8—Continued

Operations	Transcript of Think-aloud Protocol	Analysis and Images
3:21:14 PM MON PA 109,82		
3:21:15 PM MON F98 60,380		
3:21:16 PM ADL F98 PA	For now I will link PA with the root (A link is formed between PA and F98.) for place to hold characters	
3:21:20 PM MON PB 260,97		
3:21:23 PM MON F98 60,380	and I'll link PD with the root... Alright,	
3:21:24 PM ADL F98 PB		(A link is formed between PB and R83)
3:21:33 PM ACT F98 PA 3	So, a couple of characters go at node PA which would be characters 3 and 6.	(Transitions for 3 and 6 are added supporting PA.)
3:21:35 PM ACT F98 PA 6		
3:21:40 PM ACT F98 PB 8	Characters 8 and 10 go at node PB. Good. Four characters, not even sweating.	(Transitions for 8 and 10 are added below PB). At the end of this phase of building, S1 makes an evaluative statement about the number of characters explained. This is the type of statement often associated with evaluations of character compatibility groups.

Table 8—Continued

Operations	Transcript of Think-aloud Protocol	Analysis and Images
3:21:41 PM ACT F98 PB 10	<p>Let's look for things that are compatible with this whole set. Just for fun.</p> <p>Um, There may not be such a set. Character 9 is compatible with character 8 and 10, character 9 is compatible with all these others. So character 9 can go with all these building still a clique so that means we put...</p>	<p>Time</p> <p>Morphological Change</p> <p>Look at matrix Look at columns</p> <p>Identify inclusion/exclusion { (9) (8,10) }</p> <p>Add to group {8,9,10}</p>

Table 8—Continued

Operations	Transcript of Think-aloud Protocol	Analysis and Images
3:23:19 PM MSL 3 1	<p>(R83 is moved to the top of the matrix)</p> <p>Wait a minute. Character 9 is not compatible with characters 8 and 10. Scratch. Big mistake. Anything else compatible? Well, I'd better reorganize this matrix so I can see what I'm doing. Uh, R83 and R85 go together.</p>	<p>Identify conflicting</p> <p>Compare columns</p> <p>Enhance patterns in matrix</p> <p>The participant makes a rapid series of moves to reorder lines in the matrix with the apparent goal of reflecting in the matrix the inferences that were made with respect to characters 3, 6, 8 and 10. The intermediate stages of this process do not appear to be used for inferences. R83 and R85 are put together at the top of the tree. R80, R82 and R84 are grouped together below R83 and R85. R81 is moved to the bottom of the matrix, apparently just to move it out of the way. (R85 is moved to below R81, apparently by accident). (R85 is moved to below R83).</p>
3:23:22 PM MSL 5 3		
3:23:29 PM MSL 3 2		

Table 8—Continued

Operations	Transcript of Think-aloud Protocol	Analysis and Images																																																																																																																																																																																																																																																																								
3:23:38 PM 3 10	MSL And R80 and Now, 80, 82, and 84.	<p>(81 is moved to the bottom).</p> <table><tr><th>SPC</th><th>1</th><th>2</th><th>3</th><th>4</th><th>5</th><th>6</th><th>7</th><th>8</th><th>9</th><th>10</th></tr><tr><td>R81</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td><td>0</td></tr><tr><td>R82</td><td>0</td><td>1</td><td>0</td><td>1</td><td>1</td><td>0</td><td>1</td><td>1</td><td>0</td><td>1</td></tr><tr><td>R83</td><td>1</td><td>0</td><td>1</td><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td></tr><tr><td>R84</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td><td>1</td><td>1</td></tr><tr><td>R85</td><td>1</td><td>0</td><td>1</td><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td></tr><tr><td>R86</td><td>0</td><td>1</td><td>0</td><td>1</td><td>1</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td></tr><tr><td>R87</td><td>1</td><td>1</td><td>0</td><td>0</td><td>1</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td></tr><tr><td>R88</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td></tr><tr><td>R89</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr><tr><td>R80</td><td>0</td><td>1</td><td>0</td><td>1</td><td>1</td><td>0</td><td>1</td><td>1</td><td>1</td><td>1</td></tr><tr><td>F98</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr></table> <p>(Matrix as it appeared prior to modification) (R80 is moved up to R82 and R84).</p> <table><tr><th>SPC</th><th>1</th><th>2</th><th>3</th><th>4</th><th>5</th><th>6</th><th>7</th><th>8</th><th>9</th><th>10</th></tr><tr><td>R83</td><td>1</td><td>0</td><td>1</td><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td></tr><tr><td>R85</td><td>1</td><td>0</td><td>1</td><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td></tr><tr><td>R80</td><td>0</td><td>1</td><td>0</td><td>1</td><td>1</td><td>0</td><td>1</td><td>1</td><td>1</td><td>1</td></tr><tr><td>R82</td><td>0</td><td>1</td><td>0</td><td>1</td><td>1</td><td>0</td><td>1</td><td>1</td><td>0</td><td>1</td></tr><tr><td>R84</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td><td>1</td><td>1</td></tr><tr><td>R86</td><td>0</td><td>1</td><td>0</td><td>1</td><td>1</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td></tr><tr><td>R87</td><td>1</td><td>1</td><td>0</td><td>0</td><td>1</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td></tr><tr><td>R88</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td></tr><tr><td>R89</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr><tr><td>R81</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td><td>0</td></tr><tr><td>F98</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr></table> <p>(Matrix as it appears after modification)</p>	SPC	1	2	3	4	5	6	7	8	9	10	R81	1	0	0	0	0	0	0	0	1	0	R82	0	1	0	1	1	0	1	1	0	1	R83	1	0	1	0	0	1	0	0	0	0	R84	1	0	0	0	0	0	0	1	1	1	R85	1	0	1	0	0	1	0	0	0	0	R86	0	1	0	1	1	0	1	0	0	0	R87	1	1	0	0	1	0	1	0	0	0	R88	0	0	0	0	1	0	1	0	0	0	R89	0	0	0	0	1	0	0	0	0	0	R80	0	1	0	1	1	0	1	1	1	1	F98	0	0	0	0	0	0	0	0	0	0	SPC	1	2	3	4	5	6	7	8	9	10	R83	1	0	1	0	0	1	0	0	0	0	R85	1	0	1	0	0	1	0	0	0	0	R80	0	1	0	1	1	0	1	1	1	1	R82	0	1	0	1	1	0	1	1	0	1	R84	1	0	0	0	0	0	0	1	1	1	R86	0	1	0	1	1	0	1	0	0	0	R87	1	1	0	0	1	0	1	0	0	0	R88	0	0	0	0	1	0	1	0	0	0	R89	0	0	0	0	1	0	0	0	0	0	R81	1	0	0	0	0	0	0	0	1	0	F98	0	0	0	0	0	0	0	0	0	0
SPC	1	2	3	4	5	6	7	8	9	10																																																																																																																																																																																																																																																																
R81	1	0	0	0	0	0	0	0	1	0																																																																																																																																																																																																																																																																
R82	0	1	0	1	1	0	1	1	0	1																																																																																																																																																																																																																																																																
R83	1	0	1	0	0	1	0	0	0	0																																																																																																																																																																																																																																																																
R84	1	0	0	0	0	0	0	1	1	1																																																																																																																																																																																																																																																																
R85	1	0	1	0	0	1	0	0	0	0																																																																																																																																																																																																																																																																
R86	0	1	0	1	1	0	1	0	0	0																																																																																																																																																																																																																																																																
R87	1	1	0	0	1	0	1	0	0	0																																																																																																																																																																																																																																																																
R88	0	0	0	0	1	0	1	0	0	0																																																																																																																																																																																																																																																																
R89	0	0	0	0	1	0	0	0	0	0																																																																																																																																																																																																																																																																
R80	0	1	0	1	1	0	1	1	1	1																																																																																																																																																																																																																																																																
F98	0	0	0	0	0	0	0	0	0	0																																																																																																																																																																																																																																																																
SPC	1	2	3	4	5	6	7	8	9	10																																																																																																																																																																																																																																																																
R83	1	0	1	0	0	1	0	0	0	0																																																																																																																																																																																																																																																																
R85	1	0	1	0	0	1	0	0	0	0																																																																																																																																																																																																																																																																
R80	0	1	0	1	1	0	1	1	1	1																																																																																																																																																																																																																																																																
R82	0	1	0	1	1	0	1	1	0	1																																																																																																																																																																																																																																																																
R84	1	0	0	0	0	0	0	1	1	1																																																																																																																																																																																																																																																																
R86	0	1	0	1	1	0	1	0	0	0																																																																																																																																																																																																																																																																
R87	1	1	0	0	1	0	1	0	0	0																																																																																																																																																																																																																																																																
R88	0	0	0	0	1	0	1	0	0	0																																																																																																																																																																																																																																																																
R89	0	0	0	0	1	0	0	0	0	0																																																																																																																																																																																																																																																																
R81	1	0	0	0	0	0	0	0	1	0																																																																																																																																																																																																																																																																
F98	0	0	0	0	0	0	0	0	0	0																																																																																																																																																																																																																																																																
3:23:41 PM 9 3	MSL What's compatible with that. OK. Nothing is compatible with that. So I have managed to explain four characters. Maybe I can do better with some other claimed set of characters.																																																																																																																																																																																																																																																																									

Table 8—Continued

Operations	Transcript of Think-aloud Protocol	Analysis and Images
3:25:48 PM MON PB 351,133	<p>(PB is moved to a location below the position where the set defined by character 5 will be placed.)</p> <p>Take a look.</p> <p>Character 1 is compatible... character 4...</p> <p>Characters 4 and 5 are compatible. Character 4 and 5 are compatible with 3 and 6 and 7 alright, I obviously started out wrong here, let's try it again.</p> <p>characters 8 and 10 are having problems, so let me try and putting characters 4, 5 and 7 together and see what we get.</p> <p>So character 5 claims that R84 lies outside the group composed of 80, 82, 86, 87, 89.R81,</p>	<p>Look at matrix Look at columns Compare columns Identify inclusion/exclusion Start new group { (1) (4) } Identify inclusion/exclusion (4,5) Add to group { (1) (4,5) } Identify inclusion/exclusion { (1) (4,5) (3,6) (7) } Add to group {4,5,7}</p> <p>Evaluate hypothesis</p> <p>Translate groups into taxa</p> <p>Here the subject is moving nodes around as the result of translating the groups into taxa. First the taxa defined by character 5 are put in place and then are ordered according to which are included within characters 4 and 7.</p>

Table 8—Continued

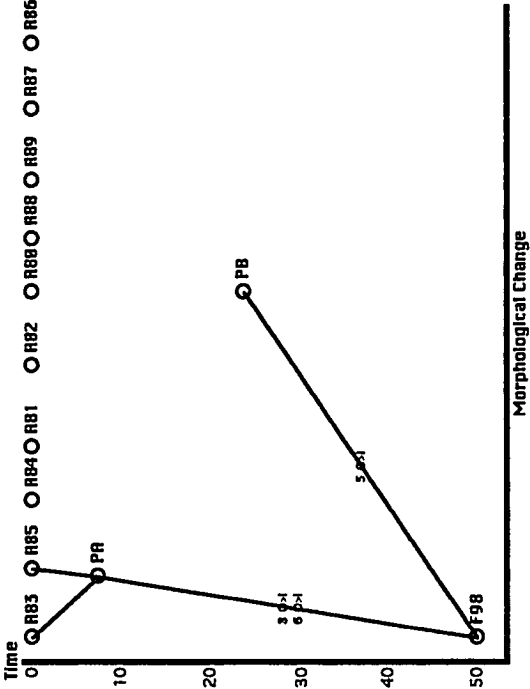
Operations	Transcript of Think-aloud Protocol	Analysis and Images
3:25:57 PM MON R81 211,30	(R81 is moved to left so as to put it outside of the group defined by character 5)	
	Alright, so, we have a clade of one, two, three, four, five, six taxa supported by character 5, correct.	Look at tree Missing taxa?
3:25:59 PM MON R82 274,30		The participant is grouping the taxa defined by character 5. (The positions of other taxa are adjusted to reduce overlapping.)
3:26:00 PM MON R80 337,30		
3:26:01 PM MON R88 391,30		
3:26:03 PM MON R86 534,30		
3:26:04 PM MON R80 338,30		
3:26:06 PM MON R80 332,30		
3:26:08 PM MON R88 375,30		
3:26:10 PM MON R86 528,30		
3:26:21 PM MON PB 333,195		(PB is moved down and to the right)
3:26:25 PM ACT F98 PB 8		(Transitions for characters 8 and 10 are removed)

Table 8—Continued

Operations	Transcript of Think-aloud Protocol	Analysis and Images
------------	------------------------------------	---------------------

3:26:26 PM ACT
F98 PB 10
3:26:27 PM ACT
F98 PB 10
3:26:28 PM ACT
F98 PB 8
3:26:31 PM ACT and I'll use PB for that killing off the (A transition for character 5 is added to support PB)
F98 PB 5

two characters. so character 5 supports a clade of 6 members,



3:26:34 PM MON
R84 169,30

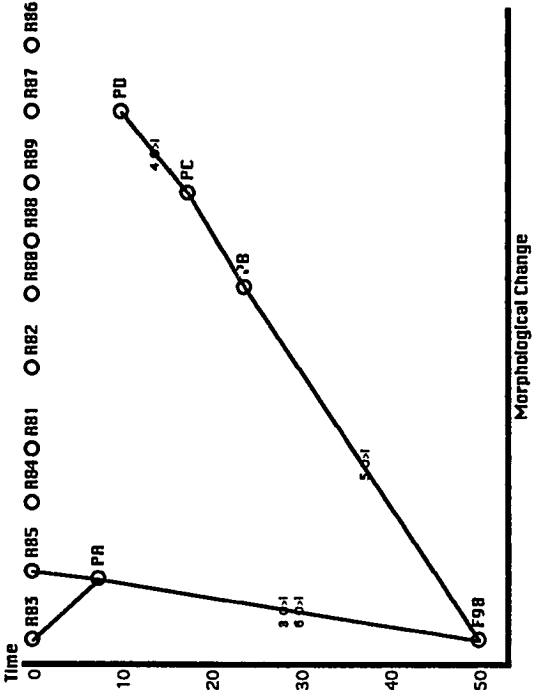
Table 8—Continued

Operations	Transcript of Think-aloud Protocol	Analysis and Images
3:26:47 PM ADN PC	character 4 goes together with that, each get a node.	(PC and PD are added.)
3:26:50 PM ADN PD		
3:26:53 PM MON PD 471,100		(PD is moved to a position roughly below R87 lined up with the link between F98 and PB)
3:26:56 PM MON PC 393,153		(PC is positioned between PB and PD)
3:26:57 PM MON PC 403,155		
3:26:59 PM MON PC 403,155		
3:27:00 PM MON PB 339,196		
3:27:04 PM MON PC 406,154		
3:27:04 PM ADL PB PC		(PB and PB are linked)
3:27:05 PM MON PD 477,99		
3:27:06 PM MON PC 412,152		
3:27:08 PM ADL PC PD		(PC and PD are linked)

Table 8—Continued

Operations	Transcript of Think-aloud Protocol	Analysis and Images
------------	------------------------------------	---------------------

3:27:14 PM ACT OK, so character 4 goes at node PD, (A transition for character 4 is added supporting PD)
PC PD 4 a clade of 80, 82, and 86



Missing taxa?

The participant is assuring himself that all of the taxa defined by character 4 are present and are organized correctly. (R88 is moved out of the way to the right) (R89 is moved out of the way to the left)

3:27:24 PM MON
R88 593, 30
3:27:27 PM MON
R89 250, 30

Table 8—Continued

Operations	Transcript of Think-aloud Protocol	Analysis and Images
3:27:57 PM MON R88 380,30		
3:27:59 PM MON R87 318,30		
3:28:00 PM MON R89 270,30		
3:28:08 PM MON R89 270,30		
3:28:09 PM MON PB 340,193		
3:28:11 PM ADL PB R89		(PB and R89 and linked together)
3:28:13 PM MON R87 318,30		
3:28:14 PM MON PC 417,151		
3:28:15 PM ADL PC R87		(PC and R87 are linked together)
3:28:39 PM ACT PB PC 7		(A transition for character 7 is added supporting PC)
	And then character 7 supports a clade of 80, 82, 86, 87, and 88 right, and then character so character 5 supports this clade. Character 7 supports this clade. Is that correct? That is not correct.	Missing taxa?
		Missing character?
		Missing character?

Table 8—Continued

Operations	Transcript of Think-aloud Protocol	Analysis and Images
3:28:44 PM R88 380,30	MON alright, what is character 7 claiming again? 82, 86, 87, 88 right, that is character 7, I'm right. And I have a clade OK, and let's put this in.	Consider parsimony Modify tree Build tree This is an example of the abbreviated form of parsimony consideration that is used for unproblematic additions to the tree.
3:28:46 PM R88 380,30	MON	
3:28:50 PM R88 380,30	MON	
3:28:51 PM PC 419,151	MON	
3:28:53 PM PC R88	ADL	(R88 and PC are linked)
3:28:54 PM R82 429,30	MON	
3:28:55 PM PD 479,93	MON	
3:28:57 PM PD R82	ADL	(R82 and PC are linked)
3:28:58 PM PD 484,91	MON	
3:29:00 PM PD R80	ADL	(R80 and PC are linked)
3:29:01 PM PD 489,90	MON	

Table 8—Continued

Operations	Transcript of Think-aloud Protocol	Analysis and Images
3:29:03 PM ADL PD R86 3:30:15 PM MON PB 253,249	<p>(R86 and PC are linked)</p> <p>(PB is moved closer to F98 to make more room on the right hand side of the diagram.</p> <p>Alright, so now I've managed to explain 5 characters without homoplasny. We're getting somewhere. So, uh, 3, 4, 5 6, And 7 are explained, leaving 1, 2, 8, 9, and 10 to explain. 8 and 10 we know have to be homoplasious, so let's see, 9, 80, 84, and 81. No that's not going to work on this claim.</p>	

Table 8—Continued

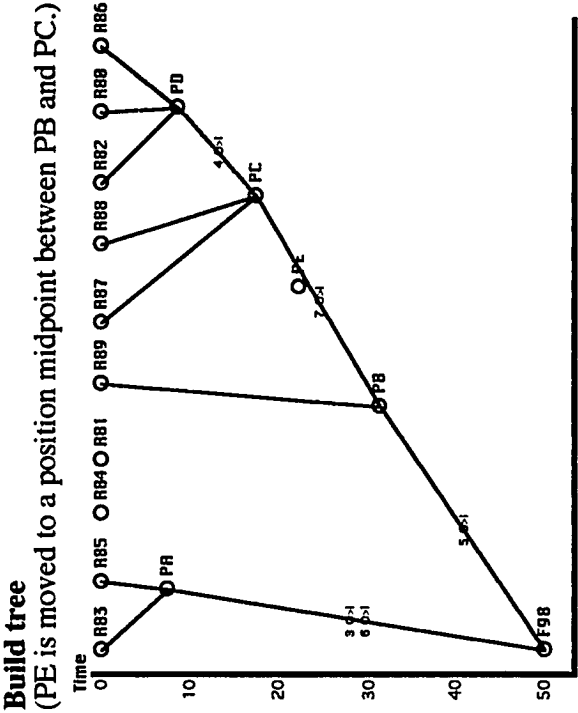
Operations	Transcript of Think-aloud Protocol	Analysis and Images
3:30:22 PM ADN PE	Uh, is character 9 compatible with character 1 and 2? character 2 is compatible with character 4 and with character 7 and character OK	(Node PE is added) Identify inclusion/exclusion { (2) (5,7,4) }
3:30:29 PM MON PE 346,186	Character 2 falls into this scheme and actually allows me to further resolve the tree. That's good I need a new node.	Add to group {5,7,2,4} Evaluate hypothesis Build tree (PE is moved to a position midpoint between PB and PC.) 

Table 8—Continued

Operations	Transcript of Think-aloud Protocol	Analysis and Images
3:31:36 PM MON R88 324,30	switch positions to make this legible.	
3:31:39 PM MON PC 419,151		
3:31:42 PM MON PE 353,185		
3:31:44 PM MON PC 419,151		
3:31:46 PM ADL PC PE	make a link	(PC and PE are linked)
3:31:49 PM ACT PC PE 2	alright.	(A transition for character 2 is added supporting PC).
3:33:40 PM ADN PF		(Node PF is added).
	<p>six characters without homoplasy and I think the rest of them are not reconcilable</p> <p>OK, we're still missing character 1, 8, 9, 10</p> <p>character 10 hasn't been looked at. it calls for R80, 82, and 84 to go together, we have no, no belief in that, I think character , characters 8 and 10, require that, character 9, 8 and 10 are identical, character 9, 1, confirmed, and 8, 10 are incompatible, mutually,</p>	<p>Consider inclusion/exclusion</p> <p>Missing characters?</p> <p>Consider parsimony</p> <p>This is the short form of parsimony -- there is no need to expressly count steps here. The claim of the character is wellevulated and the supposition that it requires homoplasy is confirmed.</p>

Table 8—Continued

Operations	Transcript of Think-aloud Protocol	Analysis and Images
3:33:43 PM MON PF 135,157	<p>Now we have to put them on here (PF is moved to a position under the unlinked taxa) somewhere.</p> <p>We also have two taxa not joined to this tree. Let's see what can be done. Character 1, links 83, 85, 84, 87</p> <p>Well, I can I assume, character 1 with a little bit of homoplasy can support another clade.</p> <p>Let's see what happens if I do it, character 1 supports, 83, 85, 84, 87, 81 all of which are free to do that except for 87, so let's make character 1 separately appear in R87 and PF</p>	

Table 8—Continued

Operations	Transcript of Think-aloud Protocol	Analysis and Images
3:34:02 PM ACT PC R87 1		(A transition for character 1 is added supporting R87)
3:34:18 PM RAL F98 PA PA PF		(F98 PA is reassigned to PA PF)
F98 PF PA		
3:34:20 PM MON PF 135,157		
3:34:23 PM MON PF 142,159		
3:34:25 PM MON PF 142,159		
3:34:27 PM MON F98 60,380		(PF is linked to F98)
3:34:28 PM ADL F98 PF		(A transition for character 1 is added supporting PF)
3:34:30 PM ACT F98 PF 1		
3:34:33 PM MON R84 169,30		
3:34:33 PM MON PF 148,158		
3:34:35 PM ADL PF R84		(R84 is linked to PF)
3:34:36 PM MON PF 148,157		

Table 8—Continued

Operations	Transcript of Think-aloud Protocol	Analysis and Images
3:34:38 PM ADL PF R81		<p>(PF is linked to R81)</p> <p>Time</p> <p>Morphological Change</p> <p>(A transition for character 9 is placed supporting R80)</p>
3:35:08 PM ACT PD R80 9	OK, minor homoplasy Fine, characters 8, 9, 10 still to work on.	

Table 8—Continued

Operations	Transcript of Think-aloud Protocol	Analysis and Images
3:35:15 PM ADN PG	Character 9 claims, R80 plus R84 plus R81 Well, I'm willing to believe in that but we can use that with a little homoplasy, to put R84 and R81 together. Node,	(Node PG is added based on support from character 9) Missing Taxa? Consider parsimony
3:35:17 PM MON PG 202,103		(PG is moved to a position below R84 and R81)
3:35:29 PM RAL PF R84 PG R84 PF PG R84		(R84 PF is reassigned to R84 PG.)
3:35:30 PM MON R81 211,30		
3:35:43 PM RAL PF R81 PG R81 PF PG R81		(R81 PF is reassigned to R81 PG.)
3:35:46 PM MON PG 202,103		
3:35:48 PM MON PF 153,159		
3:35:52 PM MON PG 181,90		
3:35:56 PM MON PG 181,90		
3:35:58 PM ADL PF PG	Combining R81 and R84... And a new link....	(PF is linked to PG).

Table 8—Continued

Operations	Transcript of Think-aloud Protocol	Analysis and Images
3:36:35 PM ACT PG R84 10		
3:36:40 PM ADN PH		(Node PH is added based on characters 8 and 10)
3:36:42 PM MON PH 440,70		(PH is moved to a position below R82 and R80)
3:36:53 PM RAL PD R82 PH R82 PD PH R82		(R82 PD is reassigned to R82 PH).)
3:37:01 PM RAL PD R80 PH R80 PD PH R80		(R80 PD is reassigned to R82 PH).
3:37:07 PM MON PH 454,67		
3:37:11 PM MON PH 454,67		
3:37:13 PM ADL PD PH		(PH is linked to PD)
3:37:16 PM ACT PD PH 8		(Transitions for characters 8 and 10 are added supporting PH).
	so with a little bit of homoplasy, characters 8 and 10 still allow R82 and R80 to go together. there.	Missing characters? Consider parsimony Build tree
3:37:18 PM ACT PD PH 10		

Table 8—Continued

Operations	Transcript of Think-aloud Protocol	Analysis and Images
3:37:22 PM MON PD 469, 107		
3:37:26 PM MON PC 403, 159		
3:37:28 PM MON PE 335, 196		
3:37:32 PM MON PD 451, 120		
3:38:50 PM MON F98 128, 380	Probably be in the new topology.... Let's see.... Alright, six characters have no homoplasy -- that we like.	<p>(Move PD down and toward F98) The nodes along the branch supported by 5 are repositioned to make more room for the transitions on PH. (Move PC toward F98) (Moves PE toward F98 (Move PD down and toward F98) (Move F98 toward middle of diagram)</p>

Table 8—Continued

Operations	Transcript of Think-aloud Protocol	Analysis and Images
	<p>Four characters have only one extra step, that's not bad.</p> <p>I can actually save two steps by putting R84 over here, but it would cost more steps than that.</p> <p>considerably more steps than that.</p> <p>Um character 1, could save a step but cost several steps to save a step well, let's see here</p> <p>character 9 again costs several steps to save a step.</p> <p>None of these homoplasies appears to help each other.</p> <p>OK, that looks like it. Now, I might as well save this</p> <p>(exporting tree)</p>	<p>Duplicated characters?</p> <p>Consider parsimony</p> <p>Postulate alternate arrangement</p> <p>Evaluate hypothesis</p> <p>Postulate alternate arrangement</p> <p>Evaluate hypothesis</p> <p>Postulate alternate arrangement</p> <p>Evaluate hypothesis</p> <p>S1 employs a heuristic: evaluating just the homoplasious characters for compatibility can indicate whether there may be another most parsimonious arrangement.</p>
3:39:55 PM CLP 5/7/95	<p>The question is are there any alternatives., Can I make a synapomorphy of the whole wad.</p> <p>I don't think so</p> <p>Character 1 has two gains and couldn't be anything else.</p> <p>Character 8 has two gains and couldn't be anything else.</p> <p>Character 10, no</p> <p>Character 9, OK</p> <p>There's no other possibility</p>	<p>(Close project)</p> <p>Postulate alternate arrangement</p> <p>Postulate alternate arrangement</p> <p>Postulate alternate arrangement</p> <p>Postulate alternate arrangement</p> <p>Postulate alternate arrangement</p>

Strategies of Phylogenetic Tree Construction

Three strategies for constructing phylogenetic trees were observed. The term "strategy" is used here to indicate more global approaches to solving problems that often can be represented as pathways through the model. The most commonly observed strategy, used by participants (1,3,4,5,6,9) was based on considering inclusion/exclusion. This was the strategy used in the example problem illustrated in Table 8. Two other strategies were consistently used only by individuals. Participant S7 constructed trees by constructing small trees, seeking duplicated taxa, and using parsimony to combine or resolve duplications. Participant S8 constructed trees by reorganizing the data matrix, considering the order of taxa in the matrix, and using a local consideration of parsimony to subdivide groups. Below I have described each strategy in more detail. In addition to a description and examples, the strengths and weaknesses of each strategy are considered. It did not seem meaningful to compare overall success among strategies because two of the three strategies were used consistently only by individuals.

Inclusion/Exclusion

The strategy of inclusion/exclusion analysis was used by the largest number of participants (1,3, 4, 5, 6, 9). This strategy is based upon partitioning a solution space into three related sets of trees. This partitioning takes place as solvers form and evaluate two kinds of hypotheses about relationships among taxa and characters: inclusion/exclusion and parsimony hypotheses.

Inclusion/exclusion hypotheses are structured by taking one character (usually the most inclusive) or a pair of identical characters and then dividing the rest of the characters into three groups based on their relationship to the first character or

characters: inclusive, exclusive, or conflicting. Inclusion/exclusion hypotheses are evaluated by comparing the numbers of inclusive/exclusive characters and incompatible characters. Parsimony hypotheses reflect branching arrangements of taxa supported by characters and are evaluated by considering the number of steps, or character state transitions, required to represent an arrangement in tree form.

For any group of taxa there is a set of possible topologies A. The number of topologies in this set increases exponentially with the number of taxa. All of the participants in the study were interested in finding subset P composed of the topologies of trees that were most parsimonious (i.e. which minimize the number of character state transitions). Participants searched for this subset by finding a related series of subsets defined by all of the character compatibility groups. For simplicity, only two subsets are described here: Subset I is the set of trees defined by all groups of inclusive/exclusive characters and subset L is the set of trees described by the largest groups of inclusive/exclusive characters. These subsets are related to one another by the rule that subset P is always a subset of I, but not always of L.

The intersection or non-intersection between sets of solutions is important for problem solving. There are 4 ways in which P and M can be related. In the case that all characters are compatible, P and M collapse into C and all three subsets are congruent (not represented). Problem 1.1 illustrates this condition. The largest inclusion/exclusion group is {1, 2, 3, 4, 5} which defines the most parsimonious tree. No other meaningful character groups are possible.

In cases of minor incompatibility, P is a subset of L which is a subset of I (Figure 72). In most of the problems (1.2, 1.3, 5.3, 2.1, 2.1a, 2.2, 2.2a, 2.3, 2.3a, 2.4, and 2.4a) P is a subset of L. In problem 1.2, the largest set of inclusive/exclusive characters is {1, 2, 4, 5}. This set also defines the most parsimonious tree. Another set

of compatible characters {3, 4, 5} is also possible, but this results in a less parsimonious arrangement of the taxa.

Problems 2.2a and 2.3a present a slightly more complicated situation. In both, the set of trees based on the largest inclusive/exclusive groups, L, derives from two competing groups of characters. In 2.2a these two groups are {4, 5, 6, 7} and {1, 3, 4, 5}. Only one of these groups results in a most parsimonious tree. In 2.3a these groups are {2, 3, 5} and {2, 3, 4}. In this case both of these groups result in most parsimonious trees.

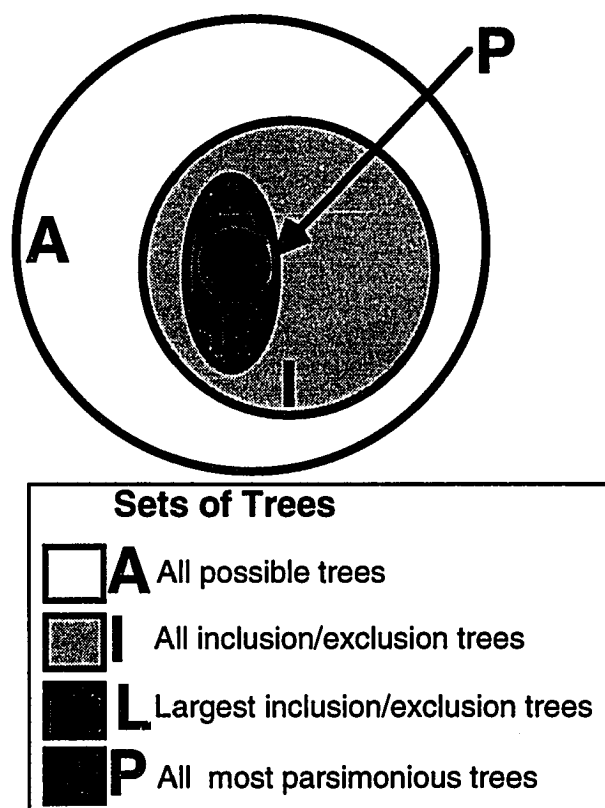


Figure 72. Most Parsimonious Trees as a Subset of Largest Inclusion/Exclusion Trees.

In cases of more significant incompatibility, P and L may only incompletely intersect (Figure 73). In these two cases, inclusion/exclusion analysis reduces the amount of the solution space which needs to be searched to find a most parsimonious solution to a problem. Two problems (1.4 and 1.5) illustrate where P and L intersect incompletely. In 1.4, the largest group of compatible characters is {1, 2, 3, 5}. The set defined by this group contains two of the most parsimonious trees, but one most parsimonious tree is defined by a smaller group: {1, 3, 5}. In this tree, both 2 and 4 are homoplasious, but the step lost in 2 is gained in 4 resulting in an equal number of character transitions.

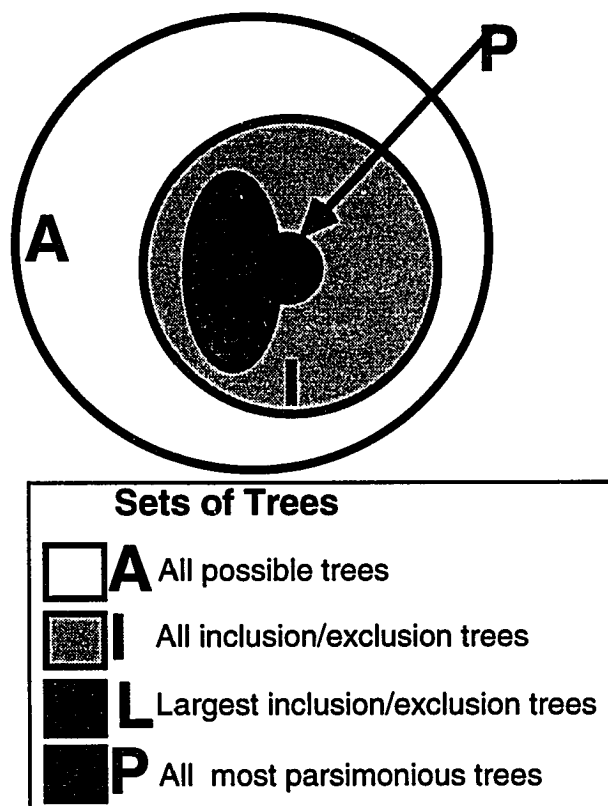


Figure 73. Incomplete Intersection of Largest Inclusion/Exclusion Trees and Most Parsimonious Trees.

In the last case, which occurs under conditions of extreme incompatibility, P and M are completely disjunct (Figure 74). In this case, seeking L will lead to incorrect solutions unless it can be counterbalanced by some other strategy to allow the problem solver to move from this set to the set of most parsimonious trees. Only one of the research problems (4.3) was of this type. In this problem, the two largest inclusion/exclusion groups are: {3, 4, 6, 8, 13, 16, 17} and {3, 6, 7, 8, 13, 14, 17}. In the most parsimonious tree, the characters 4 and 16 from the first group are false and characters 7 and 14 from the second group are false. The most parsimonious tree is based on a smaller character compatibility group composed of {3, 6, 8, 10, 13, 17}.

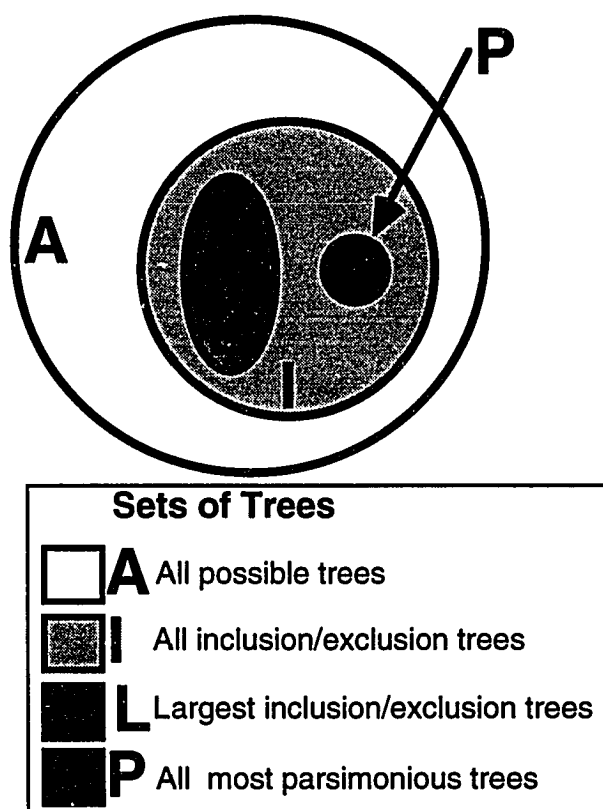


Figure 74. Non Intersection of Largest Inclusion/Exclusion and Most Parsimonious Trees.

Strengths of Inclusion/Exclusion. Inclusion/exclusion worked well across all problem types and it was the only strategy that was successful at solving problem 4.3. Inclusion/exclusion provides a mechanism for globally evaluating and comparing potential solutions without constructing them. This ability is particularly important for finding multiple topologies.

Weaknesses of Inclusion/Exclusion. Inclusion/exclusion seems to require well-developed pattern recognition skills, well-organized knowledge structures, and substantial working memory. Pattern recognition is needed to find compatible characters and knowledge structures are needed to hold the pattern of characters in memory until a hypothesis can be evaluated. The need for working memory to evaluate global hypotheses can be reduced by constructing those parts which are likely to be features of every most parsimonious tree. This provides some benefits of display based reasoning, but can result in initial errors, as in the example provided of S1 solving 5.3. Another potential means of reducing the need for working memory would be to move characters in the data matrix. If participants could move the characters in the data matrix to reflect the organization of an inclusion/exclusion hypothesis, it could facilitate display-based reasoning.

Duplicated Taxa

This strategy was used only by S7. S7 employed this strategy by converting each character systematically into a drawing. As each new character was considered, if the taxa it defined were duplicated in a small tree already drawn, the two characters were joined on a combination tree. Once all the characters were translated, any other combinations that could be made unproblematically, were made. Often this resulted in two combination trees which were exclusive or which overlapped with respect to a

single taxon. Any incompatibilities were then resolved and the result was a very parsimonious tree, usually one of the most parsimonious trees.

In appearance, it is quite different from inclusion/exclusion, but conceptually it operates on very similar principles. The result of combining all the duplicated taxa that can be combined without conflict is often the same as the tree that would result from one of the largest inclusion/exclusion hypotheses. However, because characters are combined opportunistically, the order of the characters in the matrix often contributes to which actual character group will end up defining the initial tree. Once an initial tree was constructed it was evaluated for parsimony. Considering parsimony was usually effective to optimize the tree and find a most parsimonious arrangement.

Table 9 illustrates S7 applying this strategy toward solving problem 5.3. This example does not include program operations, because this participant did not perform any problem solving activities using PI. Components from the model are in boldface and analysis of activity is in plain type. The transcript includes comments made by the researcher in parentheses. Figure 75 contains a reproduction of S7's paper and pencil phylogenetic tree.

Strengths of Duplicated Taxa. The duplicated taxa strategy makes great use of features of display based reasoning to free working memory. Each inference is reflected in the diagram immediately and each step is largely independent of all other steps. Almost all decisions are local parsimony decisions which make efficient use of working memory. This strategy results in a very efficient path to a good initial solution that can be evaluated and optimized using parsimony.

Weaknesses of Duplicated Taxa. The duplicated taxa strategy involves making combinations of duplicated taxa without considering the order in which characters are combined. Sometimes these combinations put together two characters

Table 9

Problem 5.3 as Solved by S7 Using Duplicated Taxa as a Strategy

Transcript of Think-aloud Protocol	Analysis and Images
OK. I'm ready, do you want to read off the... How many taxa?	Having observed several problems in which the participant systematically evaluated each character, I had volunteered to read off the characters from the matrix to facilitate the process and reduce clerical errors.
(We have 10 taxa. The first character has 81, 83, 84, 85, 87. Character 2 has 82, 86, 87, and 80. Character 3 has 83, 85.)	Missing taxa? Missing taxa? Missing taxa?
Aha! Hahaha. Aha! You slipped up and put something that agreed in there. OK. Oh wait, I've got to write the numbers down. 81, 84, and 87. OK.	Each of these involves looking at the matrix and constructing a small tree for the missing taxa. Duplicated taxa?
(OK, character 4 has 82, 86, and 80)	Upon recognizing that 83 and 85 are duplicates of taxa claimed by character 1, a new combination tree is drawn supported by character 1 with 83 and 85 separated out based on character 3. The initial examples of duplicated taxa lead to abbreviated considerations of parsimony. In cases where a set of taxa can be unproblematically combined, they are. Missing taxa?
So, we can do 82, 86, 80 which is character 3 and then we have 82, 86, and 87 is here and that was character... OK.	A small tree is constructed for character 4 Duplicated taxa? S7 misspeaks here -- she is actually referring to character 4.
(OK character, where are we now, character 4, no that was character 4. OK character 5, character 5 has 82, 86, 87, 88, 89, and 80.)	The subtree for character 2 is combined with the subtree for character 4.

Table 9—Continued

Transcript of Think-aloud Protocol	Analysis and Images
<p>82, 86, 87, 88,</p> <p>(89 and 80) So 1,2,3,4, this is 5. (6 has 83, 85 and that's all) 83 and 85, 6, 83, 85, OK</p>	<p>Missing taxa?</p> <p>A small tree is constructed for character 5.</p> <p>Duplicated taxa?</p>
<p>(7 has 82, 86, 87, 88, and 80) Aha! So, what is that that was 6? (that was character 7) Right, 6 is up there. So the one that was missing there, the new one was (that has 82, character 7 has 82, 86, 87, 88, and 80) OK, (And now we're at character 8? Character 8 has 82, 84, and 80). 82, Kay.</p>	<p>83 and 85 are claimed by character 3, so 6 is simply added to the 1, 3-combination tree at the location of 3.</p> <p>Duplicated taxa?</p> <p>Taxon 88 and character 7 are added to the 2,4-combination tree below character 2.</p> <p>Missing taxa?</p>
<p>(character 9 has 81, 84, and 80. And character 10 has 82, 84, and 80).</p>	<p>A small tree is constructed for character 8.</p> <p>Missing taxa? Missing taxa?</p>
<p>OK, so let's see. 80, 82, 86, 87, 88, Aha! So we can add 89 here and put 5 into that. So now, how many taxa do we have? (10) 1,2,3,4,5,6, so we've got all but 4 on a basic structure over here, so um, we can look at this, 83, is one of the ones we don't have, 85 we don't have, 81 we don't have, 84, we don't have,</p>	<p>A small tree is constructed for characters 9 and 10.</p> <p>Duplicated taxa?</p> <p>Taxon 89 and character 5 are added to the 7,2,4-combination tree below character 7.</p> <p>Missing taxa?</p>

Table 9—Continued

Transcript of Think-aloud Protocol	Analysis and Images
<p>so the only we have the conflicts between these two is 87, so my tendency at this point would be to 82, 86, 80,... 87, sometimes my mind goes faster than my hands. Fortunately that happens, I'd hate for it to be the other way around. 7, 89, 5, [...]</p> <p>So we got that one. Now we're going to combine this one with that one. We have 83, 85, and they have 3 and 6. And then we have, I know we're going to have a conflict here. I'll work it out in a minute. 81, 84, 87, 1, OK, so we got that one. So now we're left with these three in this structure and I have, the conflict was 87 at this point. 87 is here and 87 is here so at this point, I'm going to prefer to have 87 up here because its got you know, 3 characters that support it vs. 1 over there, so I would put a parallelism for 1 on 87 there and take it off here. Or we could take it off, no its not low enough down. If it were lower down you could have a reversal, but you can't do that.</p>	<p>Duplicated taxa?</p> <p>The two combination trees (1,3,6 and 5,7,2,4) are combined together. 87 is placed on both sides</p>
<p>OK, now I can look at these three. OK so this one is 82, 84, 80. 84, 82, so 82 and 84 could be placed together with this in parallel with 80, no whoops, 80 and 82, 84.</p>	<p>Duplicated taxa?</p> <p>Postulate alternate arrangement</p> <p>This is an example of duplicated taxa in the more problematic form of resolving conflicts. In this case, parsimony is considered and the most parsimonious hypothesis is selected. Subsequently, an alternate arrangement is postulated, but upon evaluation, is rejected.</p>
<p>(change tapes)</p>	<p>Missing characters?</p> <p>Character 8 is used to group 80 and 82 with a homoplasious gain in 84</p>

Table 9—Continued

Transcript of Think-aloud Protocol	Analysis and Images
<p>So we got that one. So 9 then becomes 81, 84, 80, so we could use 9 to put 81 and 4 together here and then 9 would also have to be in parallel over here, 80, but once again there's 4 characters there, so its more parsimonious to have it over here. so that's 9. 10 is 82, 84, 80. 82 and 80 are already a group, so 10 could also be in parallel there. 82, 80,... 84 over here would be 10, like that. so my guess on this one is that that's the best, the most parsimonious solution.</p>	<p>Missing characters?</p> <p>Character 9 is used to put 81 and 84 together with a homoplasious gain in 80.</p> <p>Missing characters</p> <p>Character 10 is added at the same places as character 8.</p>

which are mutually inclusive, but of which only one is consistent with the most parsimonious hypothesis. This can result in solutions which are not parsimonious.

The duplicated taxa strategy does not facilitate finding multiple topologies. The advantage of using local parsimony decisions is sometimes counterbalanced by the disadvantage of having no mechanism other than parsimony to provide a more global perspective of the problem.

Order of Divergence

This strategy was used as a primary strategy only by S8, but several participants (3, 6, 9) used it to a lesser extent to enhance inclusion/exclusion patterns in the matrix. S1 and S5 used inclusion/exclusion to enhance patterns in the matrix without considering the order of divergence.

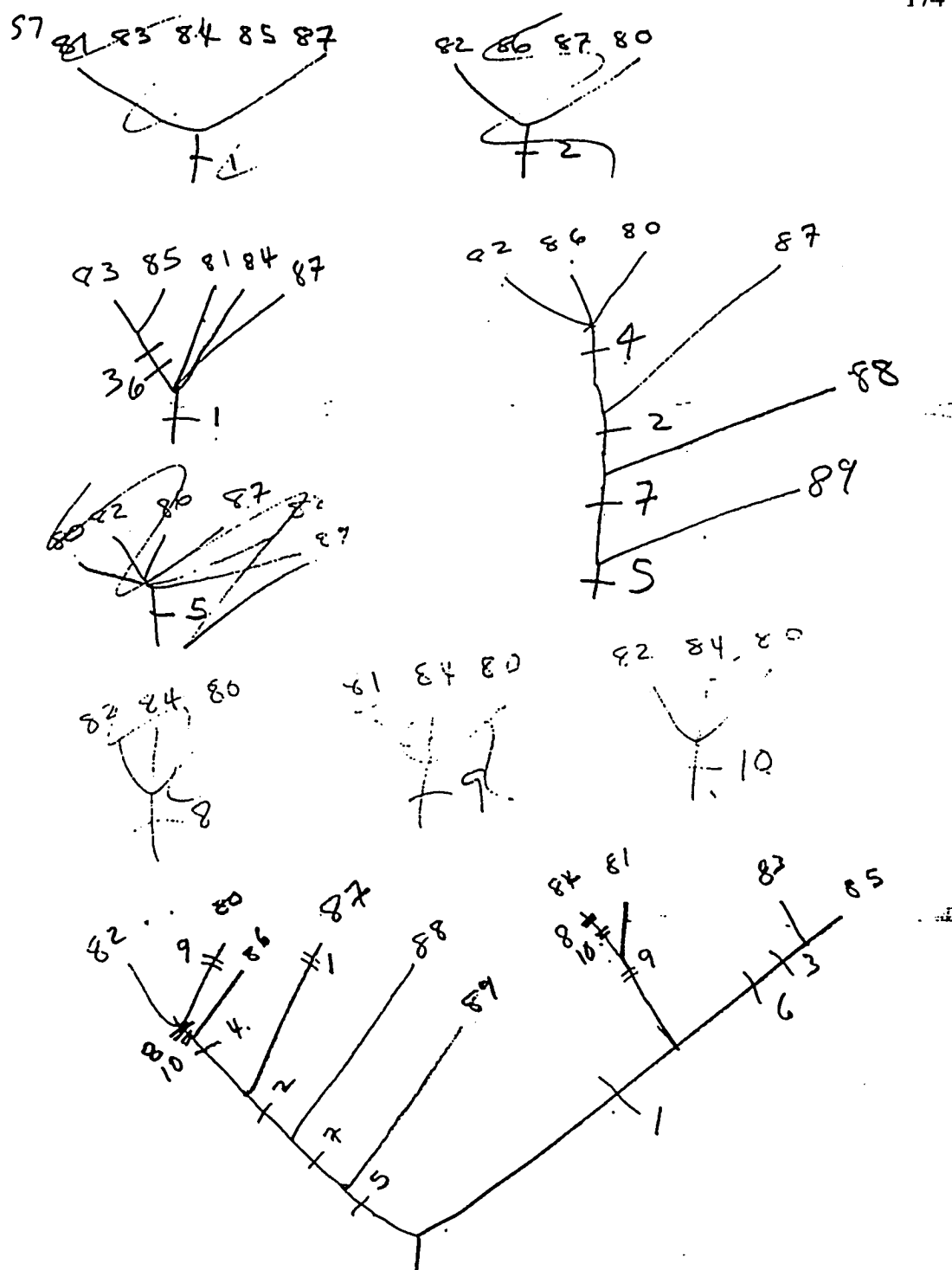


Figure 75. Reproduction of S7's Paper-and-Pencil Initial Trees and Final Phylogenetic Tree for Problem 5.3.

Participant S8 started by counting the number of apomorphies for each taxon. This was used to sort the matrix with taxa having the fewest number of apomorphies at the bottom and the most at the top. Identical and similar taxa were recognized and grouped. The matrix was sometimes sorted further to group 1's and 0's. The order of the taxa in the matrix often served as an important heuristic for causing side branches to be constructed from a main lineage. This method was often enough to allow S8 to construct one most parsimonious hypothesis. Table 10 illustrates an example of S8 using the order to divergence to construct a tree for problem 5.3. S8 used PI only to reorder the taxa and so the program operations are not presented in this example. Components from the model are in boldface and analysis of activity is in plain type. Figure 76 contains a reconstruction of S8's phylogenetic tree.

The tree that S8 constructed for this problem is not the most parsimonious arrangement of the taxa. In addition, the tree does not adequately account for the data. Taxon R84 should have apomorphies for characters 8 and 10. The fact that the recent taxa are not lined up along the top is a reflection of locations of the taxa on S8's paper and pencil tree.

Strengths of Order of Divergence. Reorganizing the matrix uses features of display based reasoning to reduce the need for working memory. Most decisions are local, rather than global decisions. The ability to organize the matrix by the number of apomorphies makes this method amenable to systematic approaches (like counting apomorphies) which can also free working memory.

Weaknesses of Order of Divergence. This strategy was not particularly successful overall and was particularly unsuccessful with difficult problems. The number of apomorphies does not correlate strongly with the order of divergence

Table 10

Problem 5.3 as Solved by S8 Considering Order as a Strategy

Transcript of Think-aloud Protocol

Analysis and Images

Aha! Here you are going both ways.
Let's see, these are, you have now 10 taxa
and, no (10 taxa and 10 characters) 10
taxa and 10 characters and two states for
each. Hmm.
Yeah, you have enough states.
I would do the same kind of thing.

Again, we'll see if, R80, OK there are
some obvious differences here.
R81 has only two apomorphies .
2 has ... 6,
I'm assuming there are some in between.
83 has ... 3 84 has ... 4. 85 has ... 3
86 has ... 4 87 has ... 4 88 has 2. 89 has 1.
OK. 80 has ... 7, what am I saying? 80 has
7.

Clearly, I would start here by moving 80
down to the bottom, like that.

And then I would move either lets see, 88,
I would move 88 next, like that.

And I would move 81. That's a little more
dubious, but there are, into that position,
waitamoment, into that position.
OK so that's a grouping. Oh that in other
words you have all of the lowest in there.

The participant is referring to the
increased numbers of characters and taxa
in this problem.

The participant observes that there are
enough characters to fully resolve the
taxa and indicates that he intends to
solve the problem using the same method
as before.

Look at rows

Count apomorphies

SPC	1	2	3	4	5	6	7	8	9	10
R81	1	0	0	0	0	0	0	0	1	0
R82	0	1	0	1	1	0	1	1	0	1
R83	1	0	1	0	0	1	0	0	0	0
R84	1	0	0	0	0	0	0	1	1	1
R85	1	0	1	0	0	1	0	0	0	0
R86	0	1	0	1	1	0	1	0	0	0
R87	1	1	0	0	1	0	1	0	0	0
R88	0	0	0	0	1	0	1	0	0	0
R89	0	0	0	0	1	0	0	0	0	0
R80	0	1	0	1	1	0	1	1	1	1
F98	0	0	0	0	0	0	0	0	0	0

Sort by apomorphies

Although the subject says "80" here, in
fact 89, which has the fewest
apomorphies is moved to the bottom.

Sort by apomorphies

88 has two apomorphies and is moved to
a position above 88.

Sort by apomorphies

81 also has two apomorphies

Table 10—Continued

Transcript of Think-aloud Protocol

Analysis and Images

And at this point, 85 and 85 and 82. 85. OK. I think I'll just move 85 ... no that's 80 something. OK.

I'm doing something wrong. 83. 87, 86, 84. 87. 80 was one of the highest.... OK. That's 80. 82,

Next, 84, 86, 87 in sequence, OK. Let's see what happens. I think I would start this tree by putting together what I'm....

Oh my God! What did I do? I just organized these very nicely on character 4, didn't I? So you get eighty.... Hehehe! That was a nice little accident, so this falls apart quite nicely,

You can get out.... OK, so first, the first thing to do is to take out 80 and 82 and 86. 80, 82, 86 on character 4. The rest have 0's in that direction.

Now of those three, they also have 2 and they also have 5 and they also have 7. Oh, did I say? 2, 5, and 7. OK. Some of these are, of course, also present in others, but that's another issue. None of them have 1. 2, 4, 5, and 7.

Sort by apomorphies

85, which also has 2 apomorphies, is placed above 81.

Sort by apomorphies

83, which has 3 apomorphies is placed above 80 (probably by accident). Then 80, which has the most apomorphies, is moved to the top.

84 is placed between 86 and 87. There is no verbalization to explain this action.

These three taxa all have 4 apomorphies,

SPC	1	2	3	4	5	6	7	8	9	10
R80	0	1	0	1	1	0	1	1	1	1
R82	0	1	0	1	1	0	1	1	0	1
R86	0	1	0	1	1	0	1	0	0	0
R84	1	0	0	0	0	0	0	1	1	1
R87	1	1	0	0	1	0	1	0	0	0
R83	1	0	1	0	0	1	0	0	0	0
R85	1	0	1	0	0	1	0	0	0	0
R81	1	0	0	0	0	0	0	0	1	0
R88	0	0	0	0	1	0	1	0	0	0
R89	0	0	0	0	1	0	0	0	0	0
F98	0	0	0	0	0	0	0	0	0	0

Consider order**Build tree**

A diagonal is constructed from lower left to upper right. A branch diverging from the bottom is constructed with a transition for character 4.

Transitions are added for 2, 5, and 7

Consider parsimony

This is the first of several remarks which indicate a concern for parsimony, but these comments are essentially never acted upon in terms of constructing a parsimony hypothesis which actually considers the length of the entire tree.

Table 10—Continued

Transcript of Think-aloud Protocol	Analysis and Images
<p>Now, 82 and 80 have.... OK... All that is 8. 80, OK. 8 has 9, 82, no wait a minute 82, 8 they also have 10. There's 9, 86 has none of those. 3, it also has 7 together with the other ones, so this is a zero. And that must be also a zero. If I did that correct. Well that takes care of those three. Wait a moment. No, that's correct. OK. Now the next group to take out are the ones that have a character, that have 1's in character 1. So, that will, then include the following, unless I'm getting myself confused. No. That will be 84, 87, 83, 85, and 81. OK.</p>	<p>Consider parsimony</p> <p>80 and 82 are separated from 86 on the basis of characters 8 and 10. 9 is autapomorphically added for 80.</p>
<p>Let's see. 84, 87. 83 and 85, that's character 3, that's unique to those two. You also have character 6. So.... 3 and 6 leads to 84, did I say? 83 and 85. 83, 85. And that 83, 85,</p>	<p>Consider order Build tree</p> <p>S8 makes a list at the side of the page. A second branch is constructed that diverges from the main branch and a transition for character 1 is added.</p>
<p>... 84, 87. OK. 84 is 8, 9, and 10. is 9. OK. 84.... So there will be a 9 on which will get us 84 and 81.</p>	<p>Missing taxa Consider parsimony</p> <p>A branch from the side branch is constructed for 83 and 85 supported by characters 3 and 6.</p> <p>Missing taxa Consider parsimony</p>
<p>And that leaves us with 87. 87 has a shitload of... at least 1 in this context. What will it have?</p>	<p>A transition for 9 is added on the side branch and 84 and 81 are placed above it.</p> <p>Missing taxa Consider parsimony</p>
<p>I'm not sure I'm doing this right, but this is general way I would do it.</p>	<p>A branch for 87 is added above the branch for 83 and 85, but below the transition for character 9. Homoplasious additions for 7, 5, and 2 are added. S8 seemed to consider his participation as describing how he solved problems, rather than attempting a rigorous application of his methods.</p>
<p>7 and 5 seem, to be, go out. I may be misreading my number. So we have enough. That one is resolved, you can resolve that one reasonably.</p>	

Table 10—Continued

Transcript of Think-aloud Protocol

Analysis and Images

That leaves us with the bottom 0's, 88 and 89 have 5 together, there is a shitload of homoplasies in this and of these 89 and 88, 88 and 89, yeah. Of course 88 and 89 are the only ones we're really working with, yeah? 88. Um, yeah. 89. 88. 7.

So that's the way I would work this tree. [...] I'm not so sure I resolved it correctly or the minimum number but that's the way I would work that particular problem.

Consider order
Consider parsimony
Build tree

A transition for 5 is added near the top of the main branch. A side branch for 88 is added. 89 is placed at the top of the branch with a transition for character 7. This is another example of S8 remarking on parsimony without actually counting steps or considering how the tree might be improved.

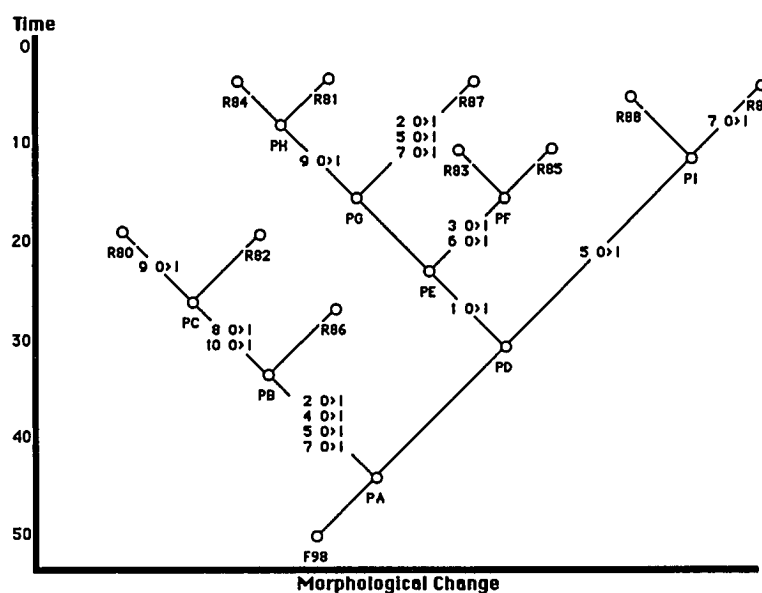


Figure 76. S8's Solution to Problem 5.3.

especially when homoplasy is common. This strategy also does not provide any mechanism for finding alternate topologies.

Seeking the Culprit

In problem 5.1, S1 said, "[There are] problems with taxon R88, a lot of incompatibility centers there." In a similar vein, while solving problem 4.3, S4 said, "OK, so [R] 89 is ... is the problem-child in all of this." Both of these participants are identifying a phenomenon which S7 had a name for:

I always look for what I call "the culprit". ... There's quite often a pattern of one or two taxa that if you eliminate those from consideration, everything falls into place.

Participant S7 employed this strategy while solving 4.3. She pulled that taxon out (R89), which greatly simplified the remaining combination trees. She then constructed a solution which was correct for the remaining 4 taxa. Afterwards she said there were several options. You could try to put the taxon in and in some cases, she said it would fit well. You could also simply set it aside and report that it was a taxon that fit poorly with the other taxa. Only the botanists (2, 7, 9) mentioned the problem of identifying hybrids. One of S7's research interests was the detection of hybrids and problematic taxa represent good candidates for further study (see Content Knowledge below). Her attempt to place "the culprit" back into problem 4.3 failed to produce a most parsimonious solution, but this strategy fits well with Sober's (1980) argument that rather than appealing to global principles, biologists should make arguments about resolving phylogenetic conflict on biological grounds.

Tree Revision Strategies

The revision problems asked participants to construct a new solution for a problem in the light of additional data. This data could either be new taxa (2.1, 2.3, 2.4) or new characters (2.2). There were two basic approaches to revision: optimize the old tree to account for the new data or simply start over from the beginning.

Placing new characters onto the old tree usually began by considering the parsimony of the character and adding transitions for the character to explain it in the most parsimonious way possible given the current structure of the tree. The number of steps on this tree then served as a baseline for evaluating optimized topologies.

Placing new taxa was more problematic. When the new taxon did not require a reorganization of the relationships among existing taxa, a simple inspection of the tree was sufficient to decide where the new taxon should be placed. There were few verbalizations associated with the decision making involved in this process. In more complex situations, participants often became confused while trying to evaluate how putting a taxon on the tree would affect global parsimony.

In revision problems there is a tension between using parsimony and going back to inclusion/exclusion to explain the new data. There are limits to participants' abilities to postulate alternate arrangements and in revision problems, it is necessary to make a decision whether to optimize the current tree using parsimony or to start over from the beginning. S1, in reflecting upon the difficulty in solving the revision problems through optimization said, (2.4):

just analyzing the optimization of characters on a tree is a difficult job, there is a bit of minor branch swapping, I can see that certain groups of characters don't change when I can't think about the other end of tree. Some local branch swapping can be simple to do, but any large-scale branch swapping, is hard to do rather than just constructing a new tree by some sort of Wagner-like process, much easier than a global branch swap.

The most successful strategy (4, 5, 7) seemed to be to add the new data onto the old tree and evaluate this tree using parsimony. If the new data appeared to fit well with old data, some minimal optimization would result in a most parsimonious solution. If the parsimony consideration indicated a poor fit, inclusion/exclusion could be used to construct a new hypothesis based on a global reevaluation of the

data. Finally, the new topology can be evaluated using the optimized old topology as a baseline.

Content Knowledge

Content knowledge played little direct role in problem solving. For participants (1, 3, 4, 5, 6, 8) content knowledge played at most tangential roles during problem solving. These participants viewed the problems as an exercise of tree building methodology. When content knowledge was brought to bear it was typically in the context of interpreting a solution of a problem and therefore did not contribute in any tangible way to the solving of the problem. In these cases, participants seemed to be trying to interpret what biological phenomenon might have given rise to a phylogenetic tree like the one they were observing, but this knowledge did not appear to affect what solutions were constructed or accepted.

Occasionally some problems reminded participants of papers they had read (1, 6) or published (1, 2, 7), or problems they had solved before (1, 4, 5, 7). S7 said while solving problem (4.3) "I did zillions of these for this paper I did years ago ... on hybridization." The use of this form of content knowledge did not appear to significantly affect how problem solving was conducted.

Referring to different character optimization schemes, several participants (4, 5, 7,) referred to ACCTRAN Accelerated Transformation and one (7) to DELTRAN (Delayed Transformation). These refer to procedures for systematically favoring the earliest possible transition of every homoplasious character (ACCTRAN) or the latest possible transition of every homoplasious character (DELTRAN). Once again, these comments were typically in the context of interpretation.

Botanists (2, 7, 9) interpreted some problem solutions as showing evidence of hybridization. For example, S7 solving 5.3 said:

I would seriously look at the parentage of [R] 84. As a botanist, I would look at one parent being [R] 81 and one parent being one of these two guys, [R] 82 and [R] 80, and I would say, what's the possibility that this is a hybrid and so I would go back and look at the plant. Where does it grow? Is it intermediate? Does it have morphology? ... and it would lead me into an investigation of the origins of taxon [R] 84.

Hybridization occurs rarely among animals and none of the zoologists mentioned hybridization.

A few participants made connections between problem-types and biological situations. When the revision problems were introduced to S7, she replied, "The missing link?" To her, this provided a context in which solving these problems made sense. S9, more than any other participant, sought to find some comparison or bridging analogy between the research problems and biological situations. When confronted with problem 2.4, in which a new taxon is added which is apomorphic for every character, S9 said, "Is that informative?" I interpreted this to mean that she questioned why anyone would conduct the analysis under those circumstances. She saw phylogenetic inference as a means of gaining insight into the relationships among taxa. Adding this taxon to the group of taxa without adding additional characters, did not seem likely to provide much insight into the situation.. She continued, saying:

It has everything, so it could go anywhere. It has everything. Well, I think I'm going to just start over. It has everything a taxon with everything. A hybrid? Taken all the gene plasm and fused it. What would that mean?

No other participant brought such a range of content-based interpretations to the problems.

Several participants (4, 7, 8, 9) mentioned how a problem solution might lead to subsequent investigation of the taxa or characters. S8, described how phylogenetic analysis could provide evidence for subsequent investigation:

the presence of a structure [should be] linked to details of the structure. That should happen. If you don't get that, if you start getting those sorted out in the wrong direction, then there is something wrong with your character set.

Related to this point, several participants (2, 6, 8, 9) mentioned the unrealistic nature of phylogenetic inference as presented in the research. S2 described the problems as "spooky" and "abstract" and indicated he felt as if he visualized the organisms and could imagine the characters changing. S6 expressed a similar pattern of reasoning: "the fact that it happens to be character 1 is immaterial, I think of things in terms of spinnerets, 2 segmented or 3 segmented." S9 made several comments about the lack of appropriate content knowledge for solving the research problems of which the following statement is typical,:

Of course, when you're actually working with real features of plants, again, certain things have more weight, for whatever reason. If you've had the ontogeny of the characters or something like that, then you've got maybe intuitive reasons giving certain characters more weight.

For these experts, the use of content knowledge may be an integral part of solving real phylogenetic problems.

Organization of Knowledge

Experts organize the knowledge they acquire about the problem during problem solving. First, experts create hypotheses that organize the data in forms that reflect inferences about relationships among the characters and taxa and which have implications for the construction of a phylogenetic tree. These hypotheses take three forms: hypotheses of order, inclusion/exclusion, and parsimony. In addition, experts organize their knowledge about phylogenetic trees through the recognition of particular classes of solutions and tree forms.

Only participant (8) predominantly solved problems through the generation of hypotheses of order, although another participant (S6) resorted to using order in difficult problems (4.3 and 5.3), and two other participants (3, 9) used order as one means of enhancing inclusion/exclusion patterns in the matrix. Generating hypotheses

of order usually involved reordering the data matrix by evaluating (or counting) the number of apomorphies each taxon contained and placing those with the most apomorphies at the top and those with the fewest at the bottom. This order, sometimes with modification, could be translated into the order taxa had diverged from the common ancestor. This method is based on a generalization which holds true only when there is little homoplasy. This generalization is also the basis of an algorithm that can be used for constructing phylogenetic trees: the Wagner algorithm (Kluge & Farris, 1969) which iteratively uses Manhattan distances to guide each step of tree construction. Both S6 and S8 indicated that if they were to successfully solve a complex problem like 4.3, they would use a process like the Wagner algorithm, but that it would require them to use a reference manual to refresh their memory on the mechanics of applying it.

The second form of data summarization hypotheses involved organizing the characters into groups based on inclusion/exclusion. These hypotheses seemed to start with a single character and then consider how many characters were inclusive/exclusive with respect to that character and how many conflicted with it. This form of knowledge organization allowed many possible arrangements of the taxa to be compared quickly. Those that were most likely to be most parsimonious—usually those based on the largest group of inclusive/exclusive characters—were translated into taxa, arranged in the drawing field, and constructed into a phylogenetic tree.

Parsimony hypotheses reflect branching arrangements of taxa supported by characters that are evaluated by considering the number of steps, or character state transitions, required to represent an arrangement in tree form. These hypotheses are constructed either in the context of changing the status of a character and evaluating the effects on the arrangement of affected taxa and characters or in the context of

changing the arrangement of the taxa and considering the effect on the underlying characters.

Before gathering data and while defining the nature of the problems, I considered how expert knowledge might be organized and constructed a series of model problems that demonstrated the subunits from which more complex phylogenetic trees could be constructed. There was little evidence that this form of knowledge organization plays an important role during problem solving. Experts do use the concepts of nested and exclusive characters, recognize characters that must be constructed convergently, and distinguish between problems with unambiguous solutions, multiple character optimizations and multiple topologies. However, with a few exceptions, they largely did not identify particular tree forms as "types" as I had predicted they might. Autapomorphies and whole in-group synapomorphies were universally recognized. Two participants (7, 8) identified situations matching the Homoplasy 1 & 2 model problem as a "three-taxon problem." For example, S7 said (1.5) "I really don't worry about that when I'm drawing the trees because that's a common three-taxon problem we have a lot in botany. Since its solvable either way, I don't worry about it."

Several other participants made statements that indicate some attention to the forms that trees take, especially with respect to trees that have a comb-like appearance versus trees that have a more symmetrical branching pattern (which is one fundamental difference between Synapomorphy 2 and 3). S3, who noted the asymmetrical pattern that resulted from 5 taxa (the number of taxa used in series 1 and 4 and the initial problems of series 2) said while solving problem 1.5, "the numbers work out more obviously with six taxa where if you have something for three and three." S6, searching for equally parsimonious solutions in problem 1.5, said, "I might have gotten way stuck on symmetrical topologies." S4, describing the

differences between two solutions of problem 2.3, said "The only way you can get anything unambiguous out of that is if you get this kind of balanced tree cause that's the only way you can have 2 here and 2 there."

Success in Problem Solving

No one found all of the most parsimonious solutions for every problem and some problems contained solutions which were found by no-one. There were not large differences of success across either problems or solvers. Table 11 lists each topology of each problem on a different line. Each line contains the number of the topology, the number of optimizations for that topology. The number for each participant shows how many optimizations were found for that topology. A zero indicates a topology that was not described. A blank space indicates problems that were not attempted by that participant. The average is calculated by summing the number of optimizations across participants, dividing by number of participants that attempted that problem and dividing that result by the number of possible optimizations. This provides a measure of overall problem difficulty, although it is not entirely valid because some participants were admittedly less concerned with finding all optimizations than others.

Variability Across Participants

Success in problem solving by participant is summarized in Table 12 as the percentage of problems in which a participant found at least 1 most parsimonious topology and the percentage of the total number of topologies that were found by each participant. The best performance in finding at least 1 most parsimonious topology (93%) indicates finding a most parsimonious solution for 14 of 15 problems.

Table 11
Number of Optimizations per Topology by Problem and Subject

Problem	Topologies	Optimizations	S1	S3	S4	S5	S6	S7	S8	S9	Average
1.1	1	1	1	1	1	1	1	1	1	0	86
1.2	1	2	2	2	2	2	2	2	1	2	93
1.3	1	2	2	2	2	2	2	2	1	1	86
	2	2	2	2	2	2	2	0	1	1	71
1.4	1	2	0	1	2	2	2	0	0	1	43
	2	2	2	0	0	0	0	0	0	0	14
	3	3	3	3	3	3	2	2	0	1	71
1.5	1	3	2	3	3	3	2	2	0	1	67
	2	2	0	0	0	0	0	0	0	0	0
	3	3	2	0	3	3	0	2	1	2	62
	4	2	2	0	0	0	0	0	0	0	14

Table 11—Continued

Problem	Topologies	Optimizations	S1	S3	S4	S5	S6	S7	S8	S9	Average
4.3	1	6	6	0	1	5	0	0	0	0	29
5.3	1	1	1	1	0	1	0	1	0	1	71
2.1	1	1	1	1	1	1	1	1	0	1	86
2.1a	1	1	1	1	1	1	1	1	0	1	86
2.2	1	2	2	1	2	2	2	2	0	0	64
2.2a	1	2	2	1	0	2	2	1	1	0	50
2.3	1	1	1	1	1	1	1	1	1	1	100
2.3a	1	1	0	0	0	0	0	0	0	0	0
2	1	1	1	1	1	0	0	1	1	0	71
2.4	1	1	1	0	0	0	0	1	1	0	43
2	1	1	1	1	0	1	1	0	0	1	57
2.4a	1	1	0	1	0	0	0	1	1	0	43

Table 12

Average Success in Percent by Participant Across Problems and Topologies

	S1	S3	S4	S5	S6	S7 ^a	S8 ^b	S9
Average across Problems	93	93	73	87	71	93	60	60
Average across Topologies	83	70	61	70	54	65	43	52

^aS7 used the duplicated taxa strategy. ^bS8 used the order of divergence strategy.

The best performance in finding the total number of topologies (83%) represents finding 19 of 23 possible topologies. All of the subjects were able to find at least one most parsimonious solution to a problem in 60% or more of the problems. Performance was universally lower across topologies than across problems. The largest different between performance across problems and performance across topologies was S7. This may be a reflection of the duplicated taxa strategy which often provides an efficient path toward a single solution, but does not provide a mechanism for evaluating global hypotheses. The range of performance is slightly larger across topologies than across problems, but the difference is small and may be explained by the fact that S8 rarely looked for alternate topologies.

S1 was perhaps the best at finding difficult solutions. S1 was one of three that solved 4.3 correctly and the only participant to find the difficult topology associated with problem 1.4 and one of the two associated with 1.5. S3 applied the inclusion/exclusion strategy systematically to solving the problems which resulted in a very consistent performance across problems. S4 is notable for having made several minor clerical errors which resulted in incorrect solutions for 5.3, 2.4, and nearly 4.3 as well. S5 used what he called a "brute force" approach to work out all likely

solutions. This resulted in one of the most consistent performances across problems. S6 seemed to do relatively well until feeling pressured for time at the end. S7 used a the duplicated taxa strategy as a systematic approach that resulted in finding a most parsimonious solution to all but one problem, but this strategy appears weak for finding multiple topologies. S8, although often able to find one most parsimonious topology, seemed less concerned with finding alternate solutions or optimizations than others and, as a consequence, did not find as many. S9 professed a lack of recent practice. A slow start, by missing the first and easiest problem, improved and resulted in a fairly steady performance.

Variability Across Problem Types

The research problems were initially constructed with the goal of studying the variability across the factors which give rise to complexity in phylogenetic tree construction. These factors were seen as increasing numbers of solutions (series 1), characters (series 4), and taxa (series 5). In addition, a series of revision problems (series 2) was posed to assess differences between model-using and model-revising problem solving. Problem 4.3, which had increased numbers of characters proved to be the most difficult problem and was correctly solved by only 3 participants. Finding additional most parsimonious solutions proved to be the most difficult aspect of problem solving and two topologies were found by no-one (topology 2 of 1.5 and topology 1 of 2.3a).

All participants (1, 3, 4, 5, 6, 7, 8, 9) used essentially the same pattern of strategies across problem types. Some heuristics were employed for solving complex problems. The heuristic by participants (4, 5) of listing character distributions and then organizing characters by inclusion/exclusion seemed particularly useful for solving problem 4.3. Several participants (1, 4, 5, 6) used considerations of order in

the matrix to enhance the ability to recognize inclusion/exclusion hypotheses more clearly when solving problems that were perceived to be more complex (4.3 and 5.3)

One result of this study is a new conceptualization of problem difficulty as the ratio between actual and potential signal-bearing characters and whether a solution is or is not constructed from the largest inclusion/exclusion character group. It quickly became apparent that the degree of homoplasy in a problem contributed to its difficulty. Initial measures of homoplasy used to evaluate problem difficulty, like the consistency index (CI), proved inadequate. Difficulty is not simply an attribute of problems: finding each topology in a problem represents a unique subproblem and it is at this level that the evaluation of difficulty must take place. The consistency index, or the ratio of tree length and numbers of characters, is the same for all equally parsimonious topologies. Statements by participants (1, 5) suggested that difficulty was a function of the signal to noise ratio. I first eliminated the characters which could contribute no signal to a topology: autapomorphies and whole-ingroup synapomorphies. I then used the remaining characters to calculate the ratio of the number of non-homoplasious characters and the total number of remaining characters. This value, or Signal Index (SI) is inversely proportional to the number of homoplasious characters in a problem.

In problem 1.5, all of the characters (5) can potentially contribute to the signal. Two of the topologies have 3 non-homoplasious characters (2, 3, and 5) yielding an SI of .6. The other two topologies have only two non-homoplasious characters (3 and 5), resulting in a lower SI of .4.

Across the problems constructed for this research, the SI ranged from 0.26 to 1.00. When the SI is plotted against the percentage of participants that found a topology (Figure 77), there appears to be a positive relationship. This relationship is described in more detail below.

Difficulty also appeared to be related to whether or not a topology was based on the largest inclusion/exclusion character group. A conceptual description of this phenomenon is provided in more detail with the description of the inclusion/exclusion strategy above. Topologies based on groups of characters other than the largest group were rarely discovered. There were three problems where this occurred: 1.4, 1.5, and 4.3. In Figure 77, the points for the topologies associated with 1.4 and 1.5 occur at .4 and .6 SI and below 20%. The point for topology tree associated with problem 4.3 is the point closest to the ordinate.

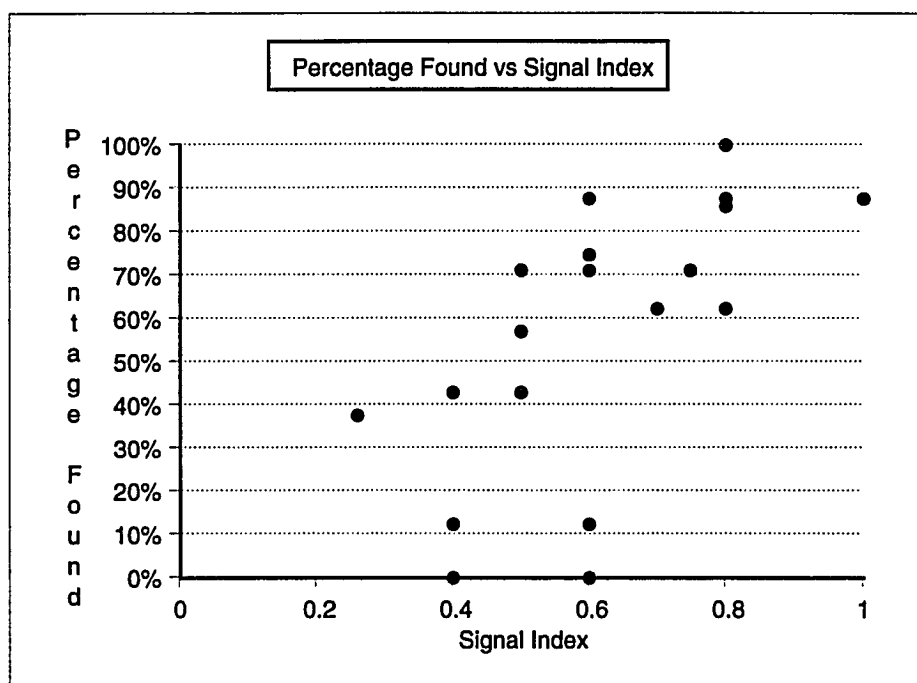


Figure 77. Percent of Participants That Found Topologies Plotted Against Signal Index.

The two factors, Signal Index and set membership (whether solutions were based on the largest inclusion/exclusion group) were analyzed as factors related to percentage of topologies found by participants using multiple regression with Minitab

(Minitab, 1992). Both factors, Signal Index ($p=.017$) and set membership ($p=.031$) were statistically significant. When both were considered, the regression function was highly significant ($p<.001$) and accounted for 59% of the variability (adjusted $R^2 = 54.7$). The regression function was $P = -0.144 + 0.376 \text{ SI} + 0.724 \text{ M}$, where P is the percentage of participants that found a topology, SI is the signal index, and M is the set membership of the topology operationalized as 0 for topologies derived from one of the largest inclusion/exclusion groups, or 1 for topologies not based on a largest inclusion/exclusion groups. One observation, identified as an outlier, resulted from the fact that one topology for problem 2.3a was found by no-one. This case is described in more detail below.

Factors that contribute to variability in performance include: the numbers of characters, taxa, and solutions; the order in which solutions were considered by a participant; practice effects; and fatigue. Most participants were able to find at least 1 most parsimonious topology for most of the problems regardless of these other factors.

The number of characters by itself would not be a good predictor of difficulty. If the characters were perfectly consistent, finding the solution would be fairly trivial regardless of the number of characters. Increasing numbers of characters in situations of homoplasy, however, magnifies the difficulty of finding a solution.

Increasing the numbers of taxa in a problem did not substantially increase its difficulty in the range used here. As the numbers of taxa increase, the problem becomes more complex and time consuming, but in the case of a set of characters that was perfectly consistent across the taxa, finding a solution would not be difficult.

Increasing the numbers of solutions in a problem by itself does not seem to affect the difficulty in finding a single most parsimonious solution, although finding all of the alternates is extremely problematic. The difficulty of finding any particular

solution is described above as a function of whether that topology is within or outside the set of trees defined by the largest inclusion/exclusion hypotheses. However, it is likely that whether a problem has multiple topologies or not and the order in which topologies are found, are factors that interact and contribute to difficulty in complex ways. The minimum length topology for Problem 4.3 was found by three of the eight (38%) participants who attempted this problem. This is substantially higher than the regression function would predict (4%). The fact that this problem has only a single solution, in spite of the fact that there is no solution within the set defined by the largest inclusion/exclusion character groups, may be a factor that contributed to participants finding the solution. Experts are often capable of determining whether a given solution represents a minimum length tree, but are typically unable to tell whether other solutions exist. This means that experts can work purposefully until they find a tree that looks like a minimum length tree. Up to that point, they are usually capable of finding ways to improve the tree. Having found a best tree, however, no expert possessed a systematic means of moving from that best tree to other best trees. Heuristics described with the model component "Postulate alternate topology" do not consistently result in pathways to alternate topologies.

It is likely that practice effects and fatigue played a role in variability of success across problem types. S9's failure to solve problem 1.1 and subsequent improved performance may be the result of practice effects. Problem 2.4 and 2.4a was always presented last and fatigue undoubtedly contributes to the poor performance in these problems. To some extent, practice effects and fatigue counterbalance each other. In any event, neither effect appears particularly large other than in the examples cited.

Errors

In phylogenetic tree construction, errors occur in four ways: when (1) not all most parsimonious trees are found, (2) no most parsimonious tree is found, (3) a less parsimonious tree is accepted as the most parsimonious tree, or (4) there is an incompatibility between the tree and the data. The first category of errors were the most common. Below, the errors are examined in more detail and when a cause for an error can be identified, it is described.

1. Not All Most Parsimonious Trees Found

Most participants were able to find one of the most parsimonious trees for almost all of the problems. No participant managed to find all of the most parsimonious solutions for every problem. Although participants (8, 9) did not place a high priority on finding other most parsimonious trees, most expressed at least some concern for finding alternate topologies. One explanation may be that participants that used the inclusion/exclusion method often found topologies within the set of trees described by the largest inclusion/exclusion character group, but rarely found trees outside of this set.

One topology from problem 2.3a that was within the set of trees based on largest inclusive/exclusive character groups was found by none of the participants. This point was identified as an outlier by the Minitab regression analysis above. In 2.3a, the set of most compatible trees results from two character groups. These groups are {2, 3, 5} and {2, 3, 4}. Both of these groups result in most parsimonious trees. The tree defined by the group of {2, 3, 5} was found by 5 of the 7 participants who attempted this problem whereas the group defined by {2, 3, 4} was found by no one.

A more careful analysis of the relationships among the characters may explain this phenomenon. As was noted in the description of model component "Identify inclusive/exclusive", participants were less likely to recognize characters as inclusive/exclusive if between them, the two characters did not account for all of the taxa. In the group {2, 3, 5}, characters 2 and 3 may be nested within 5 and then character 1, with homoplasy, can be used to group the other two taxa. In the group {2, 3, 4}, characters 3 and 4 are exclusive from one another (but between them do not account for all of the taxa) and 2 can be nested within 4. This hypothesis also uses character 1, in this case to exclude R84 and R81, with homoplasy in R86. It seems probable that participants did not recognize this inclusion/exclusion hypothesis due to the fact that since characters 3 and 4 do not account for all of the taxa, their status as inclusive/exclusive characters was evaluated as questionable by most participants.

2. No Most Parsimonious Tree Found

In several cases, participants attempted problems and quit without completing a solution. Participants (2, 6, 8, 9) failed to complete problem 4.3. Table 5, in Chapter III lists which problems were attempted and completed by each participant. Excluding S2, only problem 4.3 was attempted but not completed. In all other cases, no most parsimonious tree was found as the result of a less parsimonious tree being accepted. The results for S2 are described separately below.

3. Less Parsimonious Trees Accepted

Clerical errors made during the initial inspection of the data matrix were the most common cause of trees that were not parsimonious. Characters that were not distributed properly could also result in erroneous treelength calculations. In addition to accepting less parsimonious trees, incorrectly distributed characters also resulted in

one case where a more parsimonious topology was view as equivalent with a less parsimonious arrangement of the taxa.

Participants varied with respect to their concern for checking their solutions for errors. S3, after checking for errors, said, "If were going into publication, I wouldn't be done with it. Do you want that degree of finished? ... I would say, without frying my brain, that would be my best estimate." Some participants (6, 8, 9) expressed an explicit lack of concern for rigorously checking for errors. For example, S8 said, "Because of this, there is a shorter tree for that, which is something I haven't worried all that much about, frankly. ... I haven't counted the length because I'm not particularly worried about it." S8 seemed to feel he was demonstrating features of his approach rather than attempting to apply it rigorously. Explaining this, he said, "I'm not sure I'm doing this right, but this is the general way I would do it."

In two cases, less parsimonious trees were accepted in series 1. In the first case, S8 became confused while trying to resolve a three-taxon problem. His hypothesis based on the order of divergence was that "[R] 81 must be the pinnacle" and had placed R81 at the end of the main lineage. In this problem, among the three taxa R81, R86, and R89, one character defines a group of R81 and R89 and another defines a group of R81 and R86. S8's prior commitment to location of R81 led him to group the other two taxa even though that arrangement was supported by no characters:

[R] 89 and [R] 86, something along those lines. ... Do I make sense? Sort of? ... That one wasn't very good. ... I'm going to start flunking here as the datasets get more difficult. This is not the sort of thing I do on a daily set. I don't sit and do data sets.

In the second case from series 1, S9 failed to solve problem 1.1 correctly. S9 appears to have grouped two taxa based on their location in the matrix and then failed to recognize later that there was no support for this grouping.

Problem 4.3 was attempted by eight participants, but only completed by five. Of these, three were correct solutions and two were not. S3's incorrect solution was based on one of the two largest inclusion/exclusion groups. He sought to evaluate his hypothesis for parsimony by seeking duplicated characters, but failed to identify the correct solution. (See the description of the duplicated characters model component to see a more detailed description of S3's search for a more parsimonious solution to 4.3).

Participant S7 also attempted a solution for problem 4.3. Using the duplicated taxa strategy, she joined all unproblematic characters into combination trees and then attempted to reconcile the differences between them. In order to find the correct solution, she would have needed to use characters {1, 9, 10} to support a grouping of R80 and R82. Characters 1 and 10 were put together in a combination tree, but character 9 was grouped with character 19 and its potential to provide support went unrecognized. Her approach to solving the problem after the duplicated taxa strategy had failed is described in more detail above, under the strategy of "Seeking the Culprit."

Problem 5.3 was attempted by 8 participants and solved correctly by 5. S4 made an initial error in inspecting the data matrix and transcribing the data which resulted in an erroneous solution. S6 was late for an appointment and, although he reached a solution, it did not seem to be well considered. While solving this problem he said, "I don't know, this may not be right. Well, actually I can say that it probably isn't right." He was one of the participants (see below) who expressed little concern for checking solutions. Participant S8 also failed to solve this problem, primarily due to becoming distracted by a seemingly felicitous arrangement of the taxa which occurred as the result of sorting the matrix by the number of apomorphies. S8's approach to this problem is analyzed in detail in the section above that describes the

order of divergence strategy. Participant S8 also expressed a lack of concern with checking solutions (see below).

Two thirds of the errors in series 2 occurred in problems where the initial problem was solved correctly, but upon revision, a less parsimonious solution was accepted. Overall, at least one most parsimonious solution was found in 89% of the initial problems while this was true of only 68% of the problems upon revision. Series 2 was always the last set of problems that a participant solved and fatigue was undoubtedly an important factor that contributed to errors in series 2. S1, after solving 2.1 said, "I'm also starting to get fatigued. I may or may not be seeing all the possibilities these days."

S1 solved 2.4 correctly, but found less parsimonious solutions for 2.4a. S1 started solving the problem upon revision by making the assumption that the new taxon should be the sister group to a previous taxon based on a single character.

S3 solved 2.2 correctly and found the most parsimonious solution for 2.2a. In addition, he found a solution for 2.2a that appeared to be equally parsimonious. In fact, this solution displayed one character incorrectly and was actually one step too long.

S4 found less parsimonious solutions for 2.2a, 2.4 and 2.4a. S4 solved 2.2 correctly, but found a less parsimonious solution for 2.2a. S4 distributed the new characters onto the old tree homoplasiously and was unable to find the more parsimonious arrangement of the taxa. In problem 2.4, S4 omitted one of the taxa from his solution. After this was pointed out, he tried to place this taxon onto the existing tree, but none of the resulting three topologies was most parsimonious. S4 addressed the problem upon revision in the same way with similar results.

S5 solved 2.3 and found one of two topologies for 2.4, but was unable to find most parsimonious solutions for either problem upon revision. In problem 2.3, he first

tried to place the taxon onto the existing tree. He then tried to improve that tree by considering how else the two homoplasious characters (1 and 5) might be distributed. Neither tree matched either of the most parsimonious solutions. In 2.4, he considered trying to place the new taxon, but decided to start over. He constructed an initial solution with 10 steps, revised this tree until it had 9 steps, but was unable to find the most parsimonious solution.

Participant S7 correctly solved problem 2.2. Upon revision, S7 mapped the new characters onto the old topology and, considering an alternate arrangement, arrived at a solution that appeared to be one step more parsimonious. After optimizing the previous topology to accommodate this new hypothesis it seemed to have the same length. Several attempts to resolve the difficulty failed to reveal the cause. One homoplasious character could be distributed more parsimoniously as a gain and a loss rather than three gains.

S8 failed to find the most parsimonious trees to problems 2.1, 2.1a or 2.2. In 2.1, he began by constructing an order of divergence hypothesis. His difficulty in 2.1 appeared to arise from clerical errors. He put R88 and R82 together on the basis of 2, 4, and 5, not recognizing at first that R89 also shared 5. He later found it difficult to reorganize his tree to accommodate R89, and simply tacked R89 onto the tree, resulting in a less parsimonious solution. Upon revision, he corrected some of the problems with the previous tree, but other incorrect parts of the tree, that he believed to be unaffected by the new characters, were incorporated into the new tree without modification.

S8 constructed an order of divergence hypothesis to build a solution to problem 2.2. His tree was quite close to the most parsimonious tree, but he appeared to get distracted by the autapomorphy for character 4. S8 said, "So... I would take off [R] 86 is characterized by 4. That's the only [taxon] that has 4, in fact." Having dealt

with character 4, he was reluctant to group taxon R86 with R80, its sister taxon in the most parsimonious solution. He went on to successfully find the correct topology for problem 2.2a, although he made a clerical error of failing to distribute one of the characters.

Participant S9 was not able to find most parsimonious solutions to 2.2 or 2.2a and, although she did solve 2.3 and 2.4, she failed to solve either problem correctly upon revision. The problem in solving problem 2.2 seems to be related to placing a taxon whose only apomorphy is in a homoplasious character. Both solutions proposed by S9 placed this taxon within larger groups and used reversals to account for this taxon's ancestral condition. S9 constructed a new solution for 2.2a, but failed to find the best solution, instead constructing a solution with 1 step too many. S9 solved problem 2.3 correctly, but upon revision tried to place the new taxon onto the old tree and was unable to optimize the resulting topology into a most parsimonious solution. S9 also solved 2.4 correctly and, upon revision, started over. Her final tree was very close to the most parsimonious tree with the only error being an unnecessary grouping of R84 and R82. This error is similar to her previous error on problem 1.1. In this case, she put four of the taxa together on the basis of character 5. The two taxa R84 and R82 were placed together plesiomorphically on the basis that they did not share character 5.

Revision problems seemed to pose special difficulties for solvers, although this difficulty may have been compounded by fatigue. Errors of revision primarily occurred when participants tried to modify existing solutions. It may be advisable to strongly consider always constructing new solutions from scratch even when a revised topology seems adequate.

4. Incompatibility Between the Tree and the Data

Incompatibility between the tree and the data were usually the result of either clerical error, where the participant misread the data, or as a result of characters being distributed incorrectly. S3 found what he believed to be two equally parsimonious solutions to problem 2.2. One of them, in fact, contained a character that was displayed incorrectly. This resulted in a tree that had the right number of steps, but which did not match the data. In another example, S7 was convinced that one of two equally parsimonious trees for problem 2.2 should have been shorter than another.

When S7 examined the correct solution to 2.2, she said

the diagram's right, it's just that the display of the character is wrong. So you'd move 3, we already talked about moving 3, we just didn't do it because I thought it was the same. So 3 goes down there and then you lose it in [R] 88 and that makes a 10 step tree. The tree was right, the optimization was wrong.

Participant S2

One participant (S2) used an approach to solving the problems that was sufficiently different that his data were not included within the overall analysis. This participant used a methodology which has been called the Groundplan Divergence Method. The first stage of this methodology was always to translate the data matrix into a "formula" for each taxon. Each character was assigned a letter of the alphabet and if a taxon had that character in the derived state, its formula would contain that letter in upper case. Lowercase letters, representing the plesiomorphic state, were usually omitted. Table 13 illustrates a transformed data matrix. Taxon names were considered superfluous and taxa were referred to only by means of the formula. Taxa with identical formulas were considered to be one taxon.

Trees constructed by S2 began with a reference to "og" for outgroup. A line or lines would then be drawn upwards a short distance and a capital letter or letters, representing one or more characters would be put in the place of a node. S2 appeared to use trial and error to construct a path that could account for all of the derived characters. Once a node in the path accounted for all of the characters in a taxon's formula, that taxon was considered to have been accounted for. Figure 78 represents a completed solution, termed a "Wagner Tree" for problem 1.1. S2 was sensitive to parsimony, and sometimes noted points where parallel development of characters had occurred. He did not construct multiple trees, however, and could not always adequately account for parsimony.

Table 13
Data for Problem 1.1 as Redescribed by S2

Taxon	Characters
1.	AB
2.	ACD
3.	ACDE
4.	AB
5.	ACDE
6.	abcde

The Groundplan Divergence Method represents a separate methodology from Hennigian argumentation. It shares many of the same assumptions and features as the tree construction methods used by the other participants, but differs in its method of

data redescription, which facilitates comparison between taxa, but impedes comparing characters; its methods of inference, which are a product of the transformed data; and its treatment of ancestors, which are seen as potentially being represented by extant taxa.

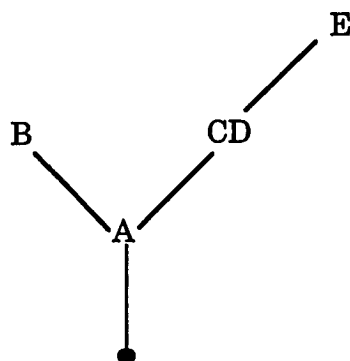


Figure 78. A Wagner Tree for Problem 1.1 as Constructed by S2.

Some parallels exist between S2 and S8, who also extensively compared rows as a means to solving phylogenetic problems. In spite of these similarities, S2 was not included with the others because the forms both of the data used to make inferences and the final representation were substantially different than those used by the other participants.

CHAPTER V

DISCUSSION

This project extends the problem-solving research tradition in science education into a new domain of biology. The results of this project corroborate previous studies in expertise and problem solving and provide new insights into how expert problem solving is conducted in biology. This research can also provide insights into the practices of phylogenetic systematics.

When this research began there was concern that, since most systematists currently use computers to find and construct phylogenetic trees, few experts in manual tree construction could be found. This proved not to be the case. Initial discussions with biologists, who reviewed early forms of the software and problems, also revealed concerns that the most recent cohort of systematics graduates, who had never practiced systematics before computers became commonplace, would not be able to engage successfully in this form of problem solving. This also proved not to be the case. This study cannot say whether these results are representative of the systematics community at large, but it can dispel the myth that the use of computer analysis consistently results in practitioners who are out-of-touch with the fundamental assumptions and practices of their discipline.

At the same time, it is clear that even relatively uncomplicated problems can quickly exceed the abilities of experts. None of the experts completely described the solution set for research problem 1.5, which had 5 taxa, 5 characters, and 4 topologies. Real problems of phylogenetic tree construction often contain dozens of taxa and characters and may have hundreds of most parsimonious solutions. As many

as 20% of studies in phylogenetic systematics rely on manual methods of tree construction (Sanderson, Baldwin, Bharathan, Campbell, von Dohlen, Ferguson, Porter, Wojciechowski, and Donoghue, 1993). Sanderson et al. (1993) report that, among the studies they surveyed, manual methods were never applied in data sets "with more than a handful of taxa" (p. 564). Based on the results of this study, I suggest that for researchers who wish to be confident that they have described or considered all of the minimum length trees, the use of a computer is indispensable.

Toward a Model of Desired Performance

The model of phylogenetic tree construction described above provides insight into how experts build phylogenetic trees. For instruction, a model of desired performance is needed that specifies the underlying forms of knowledge and thought processes that lead to good problem solving. This model should transform the most powerful aspects of the descriptive model of expert performance into a prescriptive procedural model that can be used as a goal for student learning.

Of the three overall tree construction strategies described, the inclusion/exclusion strategy offers the best potential to serve as the starting point for a model of desired performance. The inclusion/exclusion strategy was the only strategy that was successful at finding multiple topologies across all problem types. In addition, the inclusion/exclusion strategy was applied successfully both in Phylogenetic Investigator and on paper. The revised version of Phylogenetic Investigator (2.0) permits the reordering of both taxa and characters in the data matrix. This ability may facilitate display-based methods for constructing inclusion/exclusion hypotheses which could substantially reduce the need for working memory necessary to implement this strategy.

The other two strategies were applied only by single individuals and only on paper. Of the two, the duplicated taxa strategy was more successful. S7 found a most parsimonious topology for 93% of the problems. The duplicated taxa strategy works well in problems of moderate complexity and sometimes results in a more efficient path to a single solution than the inclusion/exclusion strategy. On the other hand, it provides no mechanism for finding multiple topologies. By using local rather than global decisions to combine characters, the duplicated taxa strategy generates one hypothesis efficiently, but does not provide an overview of all of the inclusion/exclusion hypotheses to guide the search for alternate topologies. Although S7 found at least 1 topology for all but one of the problems, she found only 65% of the possible topologies.

The order strategy can be useful, especially in conjunction with other strategies, but it can also be misleading. S8, who was the only participant to use the order strategy consistently, found most parsimonious topologies for only 60% of the problems and found only 43% of the possible topologies. The number of apomorphies per taxon is only a good indicator of the order of divergence when homoplasy is low or when considered systematically (e.g. as part of an algorithm like the Wagner algorithm). This method also results in a single solution and offers no guidance in seeking alternate topologies.

Based on the inclusion/exclusion strategy, Figure 79 presents a prescriptive series of procedures as an initial model of desired performance for phylogenetic tree construction. The procedures synthesize components from the model of expert performance into a set of steps that can be incorporated into teaching practice. This model should be able to produce good problem solving in phylogenetic tree construction and should be applicable to approaches using either Phylogenetic Investigator or paper and pencil.

- | | |
|------|---|
| (1) | Organize the characters, mentally, in the matrix, or on paper, to find the largest inclusion/exclusion character group hypothesis or hypotheses. (Consider the order of the taxa in the matrix as a mechanism of enhancing inclusion/exclusion patterns). |
| (2) | Translate a hypothesis into taxa by organizing the taxa in the drawing field. |
| (3) | Postulate an ancestor for each character or group of identical characters in the inclusion/exclusion hypothesis. |
| (4) | Link the most inclusive ancestors, to the next less inclusive ancestors, and continue until reaching the terminal taxa. |
| (5) | Distribute homoplasious characters. |
| (6) | Considering each homoplasious character, starting with the character that requires the most steps: |
| (6a) | Evaluate whether options that improve its distribution always result in matching losses in other characters |
| (6b) | Evaluate whether other homoplasious characters have similar distributions that reinforce each other |
| (7) | Construct other topologies based on additional inclusion/exclusion hypotheses from (1) or parsimony hypothesis from (6a) or (6b) |
| (8) | For each topology consider alternate optimizations for each homoplasious character. |

Figure 79. A Model of Desired Performance for Phylogenetic Tree Construction.

Theoretical Significance

Problem Solving and Expertise

This study corroborates previous findings in the study of expert problem solving and extends their significance within the domain of biology. Experts organize and represent information in ways that make efficient use of working memory. Experts also possess strategies and heuristics that contribute to expert performance in problem solving.

Chase and Simon (1973) proposed that a major difference accounting for expert performance was the ability of experts to encode the data into aggregated subsets which they termed "chunks." The ability to work with data in chunks rather than individual pieces permits more efficient information processing. Experts in phylogenetic inference appear to use chunks in several contexts. Inclusion/exclusion hypotheses are chunks composed of characters organized to reflect inferences about their interrelationships. Inclusive, exclusive, and conflicting characters are organized and the relative quantities of each are used for evaluation. The best inclusion/exclusion character groups are translated into taxa, to guide tree construction. Similarly, the recognition of similar and identical taxa and reorganization of the data matrix by number of apomorphies can be seen as chunking the taxa together into groups which are used to guide tree construction.

Larkin (1983) described the differences between novice and expert problem representations. In Larkin's study, novice representations tended to be constructed of objects found in the real world. Expert representations, on the other hand, not only possess objects that existed in the real world, but theoretical entities like force and momentum. Phylogenetic representations, as constructed by experts, contain both objects that exist in the real world (fossil and recent taxa) and theoretical entities like postulated ancestors and character transitions. Novice representations in phylogenetic inference have not yet been systematically examined, but anecdotal evidence suggests that novices tend to simply link existing taxa together without postulating ancestors or character transitions.

Several aspects of phylogenetic tree construction closely match Larkin's (1989) characterization of display-based problem solving: "the central premise of display-based problem solving is that the external display is the main representation of the current problem state. The phylogenetic tree during the construction process

can be used to store individual inferences and to remind the solver which parts have been considered and which he has yet to explain. The efforts of problem solvers to reorganize the matrix reflect a display-based strategy to capture inferences both in the tree and matrix.

Newell (1980) described classes of general problem-solving strategies and argued that the most powerful problem-solving strategies were domain-specific. Expert strategies for constructing phylogenetic trees do not perfectly fit any of the general patterns described. Each strategy used an approach similar to subgoal decomposition to break the problem into parts that could be solved using parsimony. The inclusion/exclusion strategy evaluated character congruence to explain the greatest number of characters possible without homoplasy and then sought to evaluate the topology using parsimony. The order of divergence strategy used an evaluation of the order in which taxa had diverged from a theoretical lineage to group the taxa and then applied parsimony to these subgroups. The duplicated taxa strategy can be considered as an implementation of the hill climbing strategy, in which local moves are performed toward a global goal for which no direct path can be discovered. S7's local moves that reduced duplication without homoplasy were employed to approach a solution and then parsimony was needed to resolve final difficulties. Parsimony was typically employed in a practice similar to the generate and test strategy, in which arrangements were postulated and evaluated.

Stewart (1989) described a series of general problem-solving heuristics and argued that these become powerful when implemented in domain specific instantiations. Table 14 lists these general heuristics and provides instantiations from phylogenetic tree construction. While some of these general heuristics do not apply to phenomena in phylogenetic tree construction as practiced in this research project, others correspond closely to patterns of reasoning observed in components of the

Table 14
Representative General Heuristics With Instantiations in Phylogenetic Tree Construction

General Heuristic	Phylogenetic Tree Construction Instantiation
1. Redescribe the data in the problem statement.	<ul style="list-style-type: none"> • Conduct an inclusion/exclusion analysis. • Reorder matrix based on numbers of apomorphies.
2. Break the problem into sub-problems.	<ul style="list-style-type: none"> • Identify competing inclusion/exclusion hypotheses and construct each in turn. • Consider alternate character optimizations before constructing new topologies. • Identify parts of problems as specific problem types (like 3-taxon problem).
3. Use external memory aids.	<ul style="list-style-type: none"> • Construct a list of character distributions to find identical and inclusive/exclusive characters. • Construct a subtree for each character.
4. Select moves that are most knowledge producing.	(Not observed/applicable)
5. Assume the most simple case first.	<ul style="list-style-type: none"> • Construct a tree based on the largest set of inclusive/exclusive characters first.
6. Consolidate your knowledge.	<ul style="list-style-type: none"> • Reorganize the matrix to group similar taxa. • Build unambiguous parts of the tree immediately.

Table 14—Continued

General Heuristic	Phylogenetic Tree Construction Instantiation
7. Check results, hypotheses.	<ul style="list-style-type: none"> • Look for homoplasious characters that seem to support one another.
8. Be systematic.	<ul style="list-style-type: none"> • Check each character in turn to consider whether it could have a different distribution.
9. Think of a related problem.	(Not observed/applicable)
10. Use knowledge to construct hypotheses.	<ul style="list-style-type: none"> • Order matrix by number of apomorphies and use that order as the basis for a first hypothesis.
11. Work backwards.	(Not observed/applicable)

model and associated strategies and heuristics. A more detailed comparison between genetics problem solving and phylogenetic tree construction is provided below that explains some of these differences.

Reif (1983) described three categories for strategic knowledge in problem solving: data redescription, solution synthesis, and solution assessment. Collins (1986) found these categories useful in her description of expert problem solving in genetics. These categories can also be applied in phylogenetic tree construction. For example, the use of inclusion/exclusion to generate a hypothesis of inclusion/exclusion character groups (data redescription), translation of groups into taxa and construction of a tree (solution synthesis), and the use of parsimony to distribute homoplasious characters and optimize the topology (solution assessment). This framework is not particularly powerful in this context, however. At a more fine-grained analysis, the distinctions between these categories blur: Is the construction of the inclusion/exclusion hypothesis "redescription" or "synthesis"? Is the parsimonious distribution of homoplasious characters "synthesis" or "assessment"? In a theoretical sense, the activities that these categories imply do occur, but it is difficult to organize the model into a form that uses these categories consistently and meaningfully.

Problem Solving in Biology

Problem-solving research in biology has focused primarily upon genetics (See Stewart & Hafner, 1994, for a review) with recent studies extending that work into the domain of meiosis (Kindfield, 1994a, 1994b). The results of this project highlight differences and similarities across domains of biology.

Kindfield's (1994) study of meiosis primarily examined "knowledge dependent representational variability and systematic uses of fine-tuned diagrams" (p. 9.) There was little evidence for representational variability as a result of knowledge

differences between participants in phylogenetic tree construction. The only example of this phenomenon is S8, who evaluated phylogenetic differences based on order of divergence. He viewed one taxon as the "pinnacle", another as most primitive, and the others as diverging from the main lineage that connected these two endpoints. The other participants constructed diagrams that were remarkably consistent with respect to form and interpretation. Understanding how experts interpret structures in phylogenetic trees, however, was not a principal question of this study and represents a potentially fruitful area of further study. To address the second point, the current study investigated only how phylogenetic diagrams are constructed, not how they are used. Future work that addresses how experts interpret and use phylogenetic diagrams may be able to draw more extensively from Kindfield's studies.

Collins (1986), Stewart (1989), and Hafner and Stewart (1995) studied genetics problem solving in the context of a particular problem-solving environment: Genetics Construction Kit (GCK) (Jungck and Calley, 1984). Model-based conceptions of problem solving have been particularly fruitful in the domain of genetics as applied to problem solving using GCK. Between genetics and tree construction there are many similarities, but also at least two important differences.

In genetics problems, as studied by Collins (1986), Stewart (1989), and Hafner and Stewart (1995), knowledge production is possible resulting in a space of potential experiments in addition to a space of potential models. In phylogenetic tree construction, as presented here, knowledge production is not possible. GCK permits the practically limitless generation of new data whereas in this research, experts were provided with a matrix of data. In practice, it is possible for additional knowledge production to occur in phylogenetic inference: a researcher can look for additional characters or taxa. In fact, researchers attempt to present the most complete data matrix possible before beginning phylogenetic analysis. The goals of phylogenetic

methods are to produce a tree that summarizes present knowledge. These results are often used to guide subsequent knowledge production, however. In this wider context, the concept of an experimental space may take on greater significance.

The tasks of genetics and phylogenetic tree construction are not parallel and genetics heuristics that relate to knowledge production and experimental spaces are not applicable to phylogenetic tree construction. The differences in the construction of problems in these two areas reflect that the study of genetics has primarily taken place within the experimental/functional tradition in biology whereas the study of phylogeny is situated within the comparative/evolutionary tradition.

A second important difference relates to the different complex of relationships among and between models, data, and activities in the two domains. Stewart and Hafner (1991) suggested that, in the context of Mendelian genetics, problem solving could be seen as using, elaborating, or revising models. Model-using problem solving encompasses the inquiry involved in the testing and justification of well structured problems that are primarily empirical in nature. Research into model-elaborating and model-revising problem solving seeks to describe how scientific inquiry addresses problems of conceptual depth and reasoning in the context of discovery.

Phylogenetic tree construction is, at one level of organization, a model-building activity. The tree that is being constructed can be considered a model that explains the data in the matrix. This model does undergo elaboration and revision during its construction and at these levels model-elaborating and model-revising problem solving can offer insight into the nature of the processes being used. Hafner and Stewart (1995) described heuristics from genetics model-revising problem solving. Some of these heuristics match phenomena observed in phylogenetic tree construction (Table 15).

Table 15

Model-Revising Heuristics With Instantiations in Phylogenetic Tree Construction

General Heuristic	Phylogenetic Tree Construction Instantiation
1. Conduct a systematic search of the initial experimental space (initial because the experimental space is reconceptualized as a result of model revision).	<ul style="list-style-type: none"> • Not applicable to tree construction. (The construction of the hypothesis being revised represents a systematic evaluation of the data prior to revision).
2. Use existing models as templates to interpret features of the cross space search that: conform to the expectations of the model; do not conform to model expectations (anomaly recognition).	<ul style="list-style-type: none"> • Construct homoplasious characters in parallel and look for matched duplicated characters at the tips of branches in order to find possible alternate topologies.
3. Use existing models as templates to postulate additional causal factors (changes in objects, states and processes) operating in a new problem type.	<ul style="list-style-type: none"> • Optimize new characters and taxa onto existing trees.
4. Utilize/build upon model revisions that have worked in the past. A model that can accommodate a large number of problem types, and thus which provides for explanatory unification within a discipline, is valued.	<ul style="list-style-type: none"> • Not applicable to tree construction. (Phylogenetic trees (i.e. the models being revised) are different from one problem to the next.)
5. Bracket the case of interest with cases on either side (interpolation).	<ul style="list-style-type: none"> • Not applicable to tree construction. (A phenomenon similar to bracketing is described in the context of model-elaborating problem-solving below).

Table 15—Continued

General Heuristic	Phylogenetic Tree Construction Instantiation
6. Break problems of increased complexity into independent subproblems, applying existing models to each. Subsequently focus upon the nature of the possible dependence between the subproblems.	<ul style="list-style-type: none"> • Remove problematic taxon, solve the problem without taxon, and then map taxon back onto finished tree. • Find all topologies for a problem and construct a consensus tree.
7. Use limiting/defining relations within existing models to: postulate additional causal factors; explore the implications of a change in one component of a model for other components of that model.	<ul style="list-style-type: none"> • Consider parsimony, e.g. compare the savings of fixing one character with the cost of breaking other affected characters.
8. Generally match a revised model to data prior to making specific mappings.	<ul style="list-style-type: none"> • Map new data onto old tree to provide a baseline before constructing different topologies.
9. Use model revisions to determine the structure of the revised experiment space, and systematically explore previously unsearched portions of that space.	<ul style="list-style-type: none"> • Not applicable to tree construction. (The concept of an experimental space does not correspond directly between genetics and tree construction).
10. Use external memory aids to represent search through both model and experimental spaces.	<ul style="list-style-type: none"> • Not applicable to tree construction. (The concept of an experimental space does not correspond directly between genetics and tree construction).

At another level, phylogenetic tree construction is a form of model-using problem solving. Experts use a model of phylogenetic inference, supported by models and assumptions, that reflects beliefs about evolutionary process. This model of phylogenetic inference results in a well-structured conception of the nature of correspondence between a data matrix and a finished tree. Problems of phylogenetic tree construction are probably incapable of presenting anomalous data that would lead to a reevaluation of the model that underlies phylogenetic inference. In this sense, the process of tree construction seems to fit the model-using problem solving category closely.

Stewart and Hafner (1991) described several observations regarding model elaborating problem solving which appear to parallel phenomena in phylogenetic tree construction. Their statement that model-elaborating problem solving may result in "Developing within model conceptual insights" appears similar to expert observations that a particular node will be true in any most parsimonious tree. Their statement that model-elaborating problem solving may result in "Linking models because they share objects, processes, or states" parallels one method of tree construction (used by S7) based on linking subtrees together based on shared taxa. Also, some participants recognized that families of potential solutions could be identified and evaluated based on common criteria. Participant S4 said, while solving problem 2.3:

This is the one. This only adds one extra step. All the other ones add 2, that the nice thing recognizing these are basically just permutations of the same thing means that basically if one of these tries adds more than that does then they're all out the window.

Regarding the nature of model-revising problem solving, Stewart and Hafner (1991) posed a series of questions. Among these they asked:

Do solvers use initial over-simplified models as starting points to develop a series of models of increasing complexity and realism? (Stewart and Hafner, 1991, p. 115)

This statement closely characterizes the nature of considering inclusion/exclusion. The solver begins by finding the largest group of completely inclusive/exclusive characters. This hypothesis is used to construct the initial hypothesis. Subsequently, homoplasious characters are mapped onto the tree and the finished hypothesis is then evaluated for parsimony. Using inclusion/exclusion results in generating an over-simplified model which is then revised, using parsimony, to create a finished tree or trees.

What anomalous data is overlooked by solvers, and is it possible to account for differences in the oversights in terms of different models that different solvers use? (Stewart and Hafner, 1991, p. 115-116)

The concept of anomalous data does not correspond exactly between genetics and phylogenetic tree construction. In genetics, anomalous data refer to differences between observed phenotypic ratios and those predicted by a model. Phylogenetic tree construction represents model construction *de novo* in response to available data. In genetics, problem solving revolves around model-using, elaborating, and revising problem solving. In phylogenetic tree construction, the primary activity can be seen as either model-using (with respect to a model of tree construction) or model-building (with respect to any particular tree being constructed). Phylogenetic data can only be anomalous in the second conception of tree construction. In this context, anomalous data can be characters or taxa that do not appear to "fit" with the others. In general, participants overlooked incompatible characters until an initial model had been constructed and then the incompatible characters were added and their effect on the structure of the tree evaluated. Several participants noted taxa that appeared incompatible with the others and one participant removed this taxon, constructed a solution, and then evaluated how the taxon could be placed back into the solution. See the strategy "Seeking the culprit:" in Chapter IV for more detail.

A second question asked by Stewart and Hafner was:

Do solvers use two models to define the extremes of a continuum of cases in which the real case is presumed to lie? (Stewart and Hafner, 1991, p. 116)

Solutions in phylogenetic tree construction are not generally organized along a continuum, although one such continuum is used to consider character optimizations: the ACCTRAN optimization constructs homoplasy as the earliest possible gain with subsequent losses and the DELTRAN optimization results in latest possible acquisition of apomorphic character states. Using these two extremes becomes important during interpretation of phylogenetic trees. Solutions related to multiple topologies are not easily conceptualized as existing along a continuum. A phenomenon similar to bracketing exists, however, in which the initial inspection does not seem to lead to a single likely solution to the problem. One way to proceed is to follow what seems to be the best path then optimize that topology to seek others. In this example, S4 has found what appears to be the best inclusion/exclusion hypothesis, but is concerned that there may be another less obvious path to a solution: S4, while solving problem 1.5 said:

I can tell with these that [...] this break is going to be at least one of the options. ... if you take a situation like this and you push all these putative homoplasies way down the tree, that's the easiest way to tell if membership in these is going to collapse and you're going to have to consider moving things around, if you push them way, way down, then you're going to remove as much support above as possible and if you've made a mistake in that initial allocation, you'll see it much more quickly ... when I see something like that, I always want to consider two possibilities: one is that these are not right and the other is the possibility that you can move all that mess way down

Another way is to identify all of the likely paths toward solutions and construct each *de novo*: In this example, S5 finds that there are two paths that seem equally likely of producing a solution:

S5 (2.4) So here its hard to pick a sort of a set of characters that think you ought to just go with because there are actually sort of 2 sets and its basically 1 and 2 vs. 4 and 5 and so what I might actually do is draw trees based on those two sets initially and see what they say about each other.

In both these examples, the solvers are using a process that is analogous to bracketing in order to estimate where solutions are and how best to ensure that all of the best solutions are found.

A third question asked by Stewart and Hafner was:

Do solvers use incomplete models as templates to capture larger or more obvious effects which can then be "factored out" to detect phenomena which would otherwise be masked or be too small to be seen. (Stewart and Hafner, 1991, p. 116)

This question closely characterizes the overall aims of phylogenetic inference. Scientists construct phylogenetic trees for the purpose of recognizing which characters can be explained most easily as the result of descent and which require some other explanation. Phylogenetic tree construction identifies those characters which closely fit a larger pattern and allows the scientist to focus research effort on the more problematic characters. Solving problem 2.3, S4 said:

What cladograms do is give you a very efficient way to curtail the range of appropriate questions in investigations. Cladograms keep you from running off and asking questions that you shouldn't be worrying about. ... the cladogram doesn't ... use the data to test the initial hypothesis, but everything that you postulate as a homoplasy on that cladogram, you've actually done a test against evidence, you say, this doesn't fit, and there are obviously two reasons for that, one is that it's real homoplasy, in which case you've got all these really interesting evolutionary things to research, or you made a mistake in which case, so what?

This aspect of model-revising, however, does not take place during phylogenetic tree construction. It is a higher level phenomenon by which phylogenetic trees are used to revise biological models in general.

Practical Significance

The practical significance of this research project is described in two contexts: improvements that contribute to the subsequent development of Phylogenetic Investigator and the development of a particular introductory biology course that uses

a problem-based approach to teaching. These descriptions are provided in concrete examples to illustrate how the results of this research become meaningful when applied in educational contexts. Problem-solving environment designers and curriculum developers should be able to apply the results of this research within other contexts as well.

Phylogenetic Investigator

This research has implications for the on-going development of Phylogenetic Investigator. PI was originally designed to allow someone to pose and solve phylogenetic problems, and draw representations for peer persuasion. This study examined its capability to facilitate posing realistic problems of tree construction, solving problems, and drawing phylogenetic trees. There was no disagreement that it was capable of presenting realistic problems. Although some experts had reservations about the biological significance of the contentless problems presented in this research, everyone agreed that these problems as represented in PI, were consistent with practices in phylogenetic biology. In terms of solving problems, PI permitted, but did not facilitate expert problem solving and most experts felt the need to supplement the environment by using paper and pencil. Below I discuss how the results of this research inform modifications to the interface for PI in terms of matrix modification and tree drawing.

Reordering data in the matrix appeared to facilitate pattern recognition and display-based reasoning. Participants (1, 3, 5, 6, 8, 9) used PI's ability to reorder rows in the data matrix and participants (1, 4, 6, 8, 9) said that the ability to move columns in the matrix would be useful as well. In addition, S7 said that she had previously engaged in this activity, although it was no longer a feature of her problem solving. One problem that was observed in moving rows was that the row selected to be

moved was not highlighted. This sometimes caused confusion about which row had been selected and whether it was actually moved to the correct new location. These insights were used to improve the data matrix in a subsequent version of PI. In the new version, both rows and columns can be moved. When they are selected, the row or column of data is outlined with a box. After a new location is selected, the column and box are moved to the new location and the box flashes to indicate the new location of the data.

The drawing environment offered several difficulties to experts. The primary complaint was that it was too cumbersome and slow. Several steps were taken to streamline its interface and improve performance. In the revised version, postulated nodes can be placed directly in desired locations rather than needing to be moved. Nodes and links can be removed using the delete key, rather than a special purpose command. Link reassignments can now be invoked more easily, either through a keystroke or by selecting a link and a node, and operate through a menu system, rather than requiring the user to enter node designations. The issue of speed was addressed by redesigning some sections of the program and by restructuring some activities and data structures. This improved the response of the program during performance critical activities, like selecting and moving items, and drawing links.

Some difficulties with the interface represent inherent differences between approaches used by experts and the environment as constructed in PI. The duplicated taxa strategy used by S7 requires multiple copies of nodes to construct single character trees. This conceptualization of the problem and its representation is simply different than that presented by PI. Experts methods and strategies are adapted to a pre-existing series of tools and representational media. Although the PI drawing environment did not facilitate expert performance on the part of most experts who used it, several participants (1, 3, 5, 6) used it successfully. Further study is required

to evaluate whether it can be used successfully to teach fundamental phylogenetic tree construction to novices.

SCI 170

This research began as an attempt to inform the development of a problem-based approach to teaching evolution. This research grew out of a search for improved problem-based curricula for the course SCI 170: Life Science for Elementary Educators at Western Michigan University. In this course, a problem-based approach has been used in genetics for several years. Below I describe how problem-based teaching has been implemented in this course using genetics and how the results of this research could inform a problem-based approach to teaching evolution.

The SCI 170 course seeks to combine the 3 P's, problem posing, problem solving, and peer persuasion approach (Stewart and Jungck, 1994) with a framework of cognitive apprenticeship (Collins, Brown, & Newman, 1989) to teach good performance in problem solving. The instructional cycle, from the standpoint of the instructor, is composed of three phases: modeling, coaching, and fading. The goal of the genetics problem solving in SCI 170 is for students to learn to apply simple-dominance, codominance, and sex-linkage models to problems; to recognize which model is appropriate in an unknown situation; and to recognize when none of the models are sufficient. (Opportunities to engage in model revision are also provided for interested students.)

Prior to instruction, students are provided with examples that illustrate phenomena that could be explained using genetics models. Genetics Construction Kit (GCK), a genetics problem-solving environment, is introduced to the students. GCK models the results of crossing fruitflies. Problems contain a simulated population of

flies which, when crossed, demonstrate the effects of an underlying genetics model. Students must reason from effects (phenotypic data) to causes (the underlying genetics model). Each time a problem is invoked, the genetics of the problem will be same, but the names associated with the phenotypic variations will be different. The genetics model operating in a problem can be concealed from the user. This requires the user to decide when a sufficient body of evidence has been generated to argue persuasively that a particular model is operating.

Initially the teacher models problem solving using GCK by applying the models in examples where the genetics is known. As problems based on different models of genetics are introduced, the instructor models the process of recognizing which model to apply. As the teacher models problem solving, appropriate strategies and heuristics are described and demonstrated. The initial problems illustrate the genetics models which students are expected to learn to apply. Subsequent problems require the application of a combination of models. After modeling the problem-solving process, the teacher solves problems with participation by students in the classroom, again suggesting places where strategies and heuristics could improve performance. Groups of students subsequently solve similar problems with the teacher acting as a coach and collaborator until students can solve examples of these problems independently.

Students perform a laboratory problem-solving assignment and are subsequently assessed. The laboratory assignment contains problems: identical with those solved as examples, that combine elements of example problems, and that present new phenomena not encountered in the example problems. Assessment usually takes the form of solving a problem similar to the example problems or answering questions about the problem-solving process. For example, a question

might ask whether or not a particular conclusion was supported by some defined subset of a body of presented data.

To be successful, instruction like that practiced in SCI 170 requires: (a) a statement that relates tasks to conceptual knowledge, (b) problems that encompass the range of phenomena to be studied, (c) models of problem solving that can lead to success; and (d) strategies and heuristics that can guide how to implement the model across the full range of situations that students may encounter. The results reported here can guide instruction of phylogenetic tree construction by: providing insight into the nature of the problems, providing a model of the problem-solving process, and identifying important strategies and heuristics.

This report addresses the relationship between the task of phylogenetic tree construction and the conceptual nature of biology from two standpoints. First, the statement of biological questions answerable by phylogenetic inference (Table 2) relates the task of phylogenetic inference specifically to the questions that it is used to answer. Second, the description of the objects, states, and processes that underlie Phylogenetic Investigator (Table 3) provides a clear statement that relates the structure of phylogenetic trees with underlying conceptual models of evolution.

The set of model problems, described in Chapter III that were developed to structure the search for expert knowledge, can be used to model phylogenetic tree construction problems to students and to illustrate fundamental relationships between the data and the tree. Insight gained from the results, reported in Chapter IV, suggests that problems can be conceptualized based on an underlying relationship between the set of most parsimonious trees and the set of solutions derived from the largest inclusion/exclusion groups. Problems can be put into classes based on whether or not all or some of the solutions to a problem are within the set of largest inclusion/exclusion groups. This conceptualization coupled with the signal index (the

ratio of the number of non-homoplasious potentially signal-bearing characters to the total number of potentially signal-bearing characters) has the potential to allow an instructor to provide a series of problems of appropriately increasing difficulty that can encompass the entire range of tree construction phenomena.

The prescriptive model of desired performance, the descriptive model of expert performance, and the descriptions of associated strategies and heuristics should all be directly applicable to the construction of curricula for teaching phylogenetic tree construction. The model of desired performance can be used as a goal for student learning and as a means of structuring a series of instructional sessions that can provide students with experiences appropriate for developing competence on each procedural component. The descriptions from the model of expert performance and associated strategies and heuristics should be useful for instructors to provide insight into how different components can be implemented and combined across a range of situations.

Questions for Future study

This study has raised a number of questions that would be appropriate for future study. These questions both clarify issues within phylogenetic tree construction and derive from topics that address the practice and application of phylogenetic inference. In addition, addressing these questions can provide further linkages between the study of problem solving in phylogenetic biology and other areas, such as genetics and meiosis.

Some of the components described in the procedural model of phylogenetic tree construction in Chapter IV could benefit from subsequent elaboration. In particular, all of the processes associated with the evaluation of parsimony could benefit from additional elaboration. For example, the component "Postulate alternate

arrangement" describes the nature of the activity that participants engaged in, but the methodology used in this study was at a course-grained and did not provide much insight into the actual processes that participants use to conduct this activity. Studies that could shed light on some of these processes and explain in more detail how they are conducted could inform how the model should be used and taught.

One realm of questions refers to interpretation of phylogenetic trees.

Understanding how phylogenetic trees are evaluated, interpreted, and used by experts could corroborate and extend much of what is proposed in this report. Several specific questions that this study raised are addressed individually below.

Do trees stand alone or must they be interpreted in the context of their data matrix? In order to evaluate trees proposed by others, do experts study the tree itself or do they recapitulate the process of construction? Most participants constructed trees based on some form of an inclusion/exclusion hypothesis and their subsequent optimization of trees was performed in the context of that hypothesis. S7 described her experience evaluating phylogenetic trees that have been proposed for publication:

I always take papers that I have to review for and the first thing I do is I take the data matrix and I go through and scratch out all the autapomorphies and I scratch out all the synapomorphies and then I start doing little character trees and combining everything together and trying to make a diagram myself and then after I get a diagram made I compare it to the one that the person is proposing as a phylogeny and usually now with the computer programs they are either shorter than the one I got or, especially if they're messy, but occasionally, I'll actually get a shorter diagram and then it's sort of interesting because then you wonder how the hell they used a computer program and got a longer diagram than you did. But with the molecular data sets it's much more difficult because of the large numbers of taxa [and characters]

Other questions for further study include: How, specifically, do experts relate structures in the tree environment to underlying models of evolution? Some comments made by experts suggest that there are different interpretations among experts used to explain the meaning and underlying processes that result in phylogenetic patterns. How do experts compare trees? Tree comparisons occur both

as the result of multiple topologies from a single data set and as the result of using different data sets. Research into how experts address these issues could complement and extend this project in significant ways.

Studies of students' initial conceptions of phylogeny and of novice performance in problem solving could also inform the teaching of phylogenetic tree construction. Meaningful learning in science occurs when students come to understand their own conceptions and their limitations and then, seek to replace those conceptions with scientific conceptions. Some conceptions held by students are resistant to instruction and understanding the conceptions held by students can provide insight into potential difficulties. Similarly, a model of novice performance can be compared with a model of desired performance to illustrate specific points that instruction could address to help students improve problem solving.

Conclusion

This research project provides insights into the nature of phylogenetic problems and problem-solving methods. This report includes a new description of the nature of phylogenetic problems and factors contributing to their difficulty that can inform the construction of problems for educational use that encompass the full range of phylogenetic phenomena. The procedural model of phylogenetic tree construction adds to the knowledge of problem-solving research and can inform desired performance for students. The descriptions of associated strategies and heuristics provide methods for implementing the model in a variety of situations. The results of this research can inform the creation of curricula that address the teaching of evolution and provide a foundation for subsequent research in the problem solving of phylogenetic inference.

Appendix A
**A Primer of Phylogenetic Assumptions,
Diagrammatic Elements, and Terms**

A Primer of Phylogenetic Assumptions, Diagrammatic Elements, and Terms

Assumptions of Phylogenetic Inference:

1. There is only one true phylogeny.
2. Shared characters are the result of homology.
3. The polarity of character states is knowable.

Elements of Phylogenetic Diagrams

Figure 1 illustrates an example phylogenetic tree created using Phylogenetic Investigator. This section describes the phylogenetic tree and its elements . Unfamiliar terms are organized alphabetically in the following section with definitions and examples that also reference this tree where possible.

The data matrix from which this diagram is generated appears in the lower right hand corner showing characters in columns and taxa in rows. The intersection between each row and column has a symbol that indicates where that taxon has the apomorphic (1) or plesiomorphic (0) form of the character.

The phylogenetic tree is constructed along two axes. The ordinate represents time divided into 50 units and the abscissa represents morphological change as a continuous, unitless variable. The small circles are nodes. Each node has a designation associated with it. Nodes that begin with "R" represent recent taxa. Nodes that begin with "F" represent fossil taxa. Nodes that begin with "P" are postulated taxa. Lines that link nodes together indicate lines of ancestor/descendant relationship. Some links contain 1 or more transitions. Each transition (e.g., "1 0>1" or "1 1>0") indicates that the referenced character (1) changed in state either from plesiomorphic

to apomorphic (0>1) or reversed from apomorphic to plesiomorphic (1>0) at some point in time along the link on which it appears.

In Figure 1, characters 1-5 are represented as being homologous. Characters 6 and 7 are homoplasious in this diagram. Character 8 is an autapomorphy and is irrelevant to the decision-making process of tree construction. An autapomorphic character is always constructed as a transition immediately prior to the taxon that possesses it.

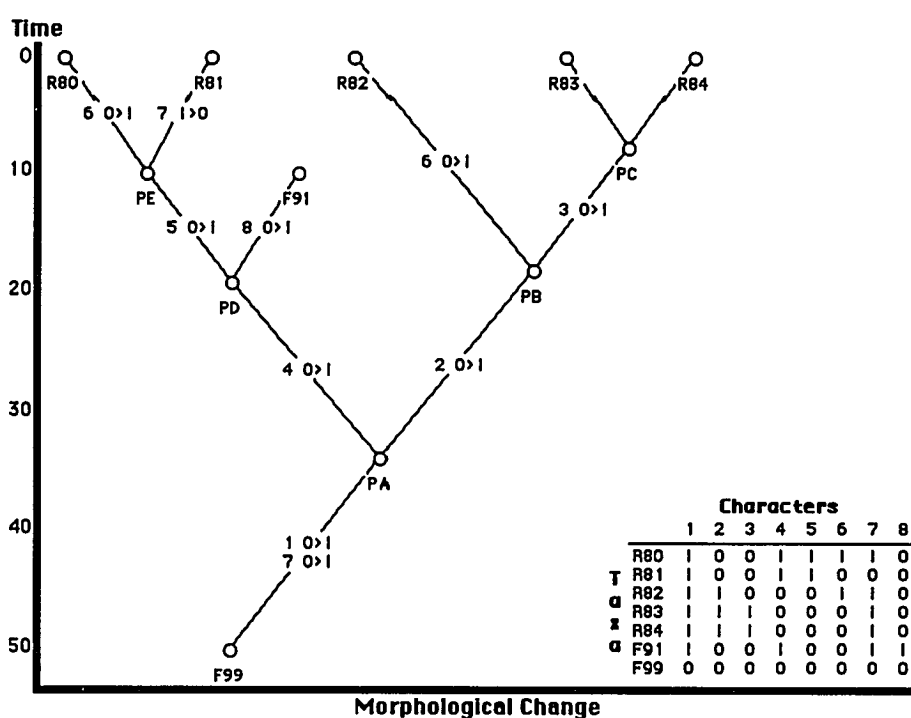


Figure 1. An Example Phylogenetic Tree

Character 1 is a whole-group synapomorphy that supports the existence of postulated ancestor PA. Character 1 is inclusive of all other characters. Character 2, which groups R82, R83 and R84, supports node PB. Character 2 is inclusive of character 3 and exclusive of character 4. Character 3, which groups R83 and R84,

supports node PC. Character 4, which groups R80, R81, and F91, supports node PD. Character 4 is inclusive of character 5 and exclusive of character 2. Character 5, which groups R80 and R1, supports node PE.

Character 6 claims that R80 and R82 are a group. For character 6 to be true, characters 2, 5, and 4 would have to be false. In other words, in order to save one step in character 6, at least 3 other steps would be required. Character 6 is most parsimoniously gained convergently in R80 and R82. Character 7 claims that all of the taxa except for R81 are a group. For character 7 to be true, characters 4 and 5 would have to be false. Saving a step in character 7 would result in at least 2 added steps. Character 7 is most parsimoniously optimized as a reversal in R81.

Terms of Phylogenetic Inference

Ancestor	A taxon, previous in time to a second taxon, from which the second taxon is descended. For example, Figure 1 proposes that a postulated taxon PC is the common ancestor of R83 and R84.
Apomorphy	An evolutionary character, usually coded as "1", that represents an evolutionarily novel state. Figure 1 proposes that character 1 is an apomorphy in all of the taxa of the ingroup.
Autapomorphy	The transition of a character that is uniquely evolutionarily novel (apomorphic) for a taxon. Character 8 in Figure 1 is an autapomorphy because it is possessed in the apomorphic state only by taxon F91.
Character	A recognizable feature that varies among taxa. For example, among ladybugs, the characters might include the presence or absence of spots. Characters are numbered, polarized, coded, and presented in columns in the data matrix (Fig. 1).
Clade	A monophyletic taxon.
Cladogram	A form of phylogenetic tree that can only show sister-group relationships. Figure 1 illustrates sister-group relationships between all of the taxa, except F99, which is claimed to be a true ancestor of all of the other taxa.
Conflict	A quality of characters that contain incompletely overlapping distributions of apomorphies. Characters 5 and 6 conflict

	because both are apomorphic for 80, but 5 is apomorphic for 81 and 6 is apomorphic for 82 (Fig. 1).
Convergence	A form of homoplasy whereby two taxa share a character that has appeared independently in separate lineages. Character 6 arises convergently in taxa R80 and R81 (Fig. 1).
Data Matrix	A summary table of states with taxa in rows and characters in columns. The data matrix appears in the lower right-hand corner (Fig. 1).
Descendant	A taxon which is the genealogical product of an earlier taxon. Taxon R84 is a descendant of PC (Fig. 1).
Exclusive	Characters whose distributions of apomorphies do not overlap. Characters 2 and 4 are exclusive of one another (Fig. 1).
Homology	The quality of characters that are shared as the result of common ancestry. See assumption 2. Characters 1, 2, 3, 4, 5 are assumed to be homologous (Fig. 1).
Homoplasy	Characters that are shared due to causes other than homology (evolutionary convergence or reversal). Character 6 is homoplasious and explained using convergence and character 7 is homoplasious and explained using reversal (Fig. 1).
Inclusive	When one character's distribution of apomorphies is a superset of another character's distribution of apomorphies. Character 2 is inclusive of character 3 (Fig. 1).
Ingroup	The group of taxa currently being studied using phylogenetic inference. In Figure 1, R80, R81, R82, R83, R84 and F91 are members of the ingroup.
Link	A line in between nodes in Phylogenetic Investigator that represents lines of ancestor/descendant relationships. The link between R83 and PC represents a hypothetical ancestor/descendant relationship between R83 and PC.
Monophyletic	A taxon that includes only the complete set of descendant taxa of an ancestral species. The group of R83 and R84 (and PC) is a monophyletic taxon (Fig. 1).
Node	A circle in Phylogenetic Investigator used to represent a taxon. R80 is a node that represents a taxon (Fig. 1).
Optimization	The process or product of distributing a homoplasious character on a phylogenetic tree. Characters 6 and 7 are optimized in Figure 1.
Outgroup	A group of taxa used to polarize the character states.

Parallelism	A convergence.
Paraphyletic	A grouping of taxa that does not reflect the underlying evolutionary relationships by removing taxa from a monophyletic taxon. A grouping of R82 and R84 is paraphyletic (Fig. 1).
Parsimony	A principle used to justify selecting the hypothesis that requires the fewest transitions and a corollary to assumption 2: By assuming homology, one also selects the hypothesis that minimizes the number of assumptions of homoplasy. Fig 80 is the most parsimonious explanation of the data.
Phylogenetic tree	A branching diagram that can illustrate both sister group and ancestor/descendant relationships among a set of taxa. Figure 1 is a phylogenetic tree.
Phylogeny	The set of ancestor/descendant relationships that form the genealogy of a set of taxa. A phylogenetic tree (Fig. 1) is a hypothetical representation of these relationships.
Plesiomorphy	A form of a character (state) which is evolutionarily preexisting for the group of taxa under study (the ingroup). Character 2 is retained in the plesiomorphic state by R80, R81, and F91 (Fig. 1). Character 7 occurs in the plesiomorphic state in taxon R81 and this is explained using a hypothesis of reversal (Fig. 1).
Polarity	Whether a form of a character (a state) is considered apomorphic (evolutionary novel) or plesiomorphic (evolutionarily preexisting). This is usually done through comparison with an outgroup.
Polyphyletic	A grouping of taxa that does not reflect the underlying evolutionary relationships by adding unrelated taxa to a monophyletic taxon. A grouping of R81, R83, and R84 would be polyphyletic (Fig. 1)
Reversal	The transition of a character that is apomorphic in some ancestor, changes polarity back to the plesiomorphic state resulting in descendant taxa which are plesiomorphic for that character. Character 7 is optimized as a reversal in taxon R81 (Fig. 1).
Sister group	The most closely related taxon to another taxon. R82 is the sister group to the taxon of R83 and R84 (Fig. 1)
State	A form of a character that is polarized as either apomorphic or plesiomorphic and coded as "1" or "0". For example, among ladybugs, the absence of spots might represent the

plesiomorphic state and the presence of spots might represent the apomorphic state.

Steps	The number of transitions required to explain a character or characters. Character 6 is explained in two steps (Fig. 1).
Synapomorphy	The transition of a character that is homologously shared in the evolutionary novel (apomorphic) condition. Character 1 is a synapomorphy for the whole ingroup (Fig. 1).
Taxon	A group of organisms that is given a name. The complete set of taxa descended from a common ancestor is a monophyletic taxon. Incomplete sets are paraphyletic and sets with extra unrelated taxa are polyphyletic. R80, R81, and F91 were it to be a monophyletic taxon because they all are hypothesized to have descended from PD (Fig. 1).
Topology	An arrangement of sister-group or ancestor/descendant relationships among a group of taxa. Figure 1 has only 1 most parsimonious topology—any rearrangement of the relationships among the taxa would require more steps than the current tree to explain all of the characters.
Transition	A point in time in a lineage at which a character is hypothesized to have changed in state. At some point between 20 and 35 units of time before the present, character 4 is hypothesized to have changed in state in taxon PD (Fig. 1).
Treelength	The steps, or number of transitions, required to explain the data matrix using a phylogenetic tree. Figure 1 requires a treelength of 10 steps to most parsimoniously explain the data in the matrix.

Appendix B
Protocol Clearance From the Human Subjects
Institutional Review Board

Human Subjects Institutional Review Board

Kalamazoo, Michigan 49008-3899
616 387 8253

WESTERN MICHIGAN UNIVERSITY

Date: March 13, 1995

To: Brewer, Steven D.

From: Richard Wright, Interim Chair

Re: HSIRB Project Number 95-03-02

This letter will serve as confirmation that your research project entitled "Expert phylogenetic tree construction using phylogenetic investigator" has been approved under the expedited category of review by the Human Subjects Institutional Review Board. The conditions and duration of this approval are specified in the Policies of Western Michigan University. You may now begin to implement the research as described in the application.

Please note that you must seek specific approval for any changes in this design. You must also seek reapproval if the project extends beyond the termination date. In addition if there are any unanticipated adverse or unanticipated events associated with the conduct of this research, you should immediately suspend the project and contact the Chair of the HSIRB for consultation.

The Board wishes you success in the pursuit of your research goals.

Approval Termination: Mar 13, 1996

xc: Hafner, Robert, SST

Appendix C
Internet Recruitment Posting

Internet Recruitment Posting

For the past couple of years Bob Hafner and I have been looking for ways to improve the teaching and learning of evolution and systematics. We have embarked on a research program to study problem-solving in phylogenetic analysis. We believe that students should be engaged in problem-solving as a means to increase not only their content knowledge, but their understanding of general and domain specific heuristics and the nature of science as an intellectual activity. We believe our work will inform development of curricula and teaching materials that can greatly improve teaching and learning of evolution and systematics. My dissertation, a study of expert phylogenetic tree construction, will be the beginning of this research program.

We have developed a computer-based problem-solving environment, called Phylogenetic Investigator (PI), that can be used for students to pose and solve problems of phylogenetic inference. This environment is different from most phylogenetic research tools in several important ways. First, PI simply provides an environment with tools where the problem solver can record and summarize data and then use that data to manually draw phylogenetic or cladistic hypotheses: It does not automatically generate trees. Second, PI can internally generate plausible data of specific problem types while varying the surface characteristics of the problems. This allows students to practice problems repeatedly and facilitates teachers being able to model the problem-solving process and to act as a coach or guide. Third, PI does not generate statistics or check the validity of trees produced by students. PI is intended for an environment where students generate hypothesis and then try to persuade peers that their approach to the problem is the best one. Toward this end, we have sought to allow students a full range of expression in this medium: students can represent

variable rates of divergence among species, anastamosing tree forms, and other patterns which are difficult or impossible to represent other programs we have examined.

Eventually, we hope to extend this work by studying other aspects of phylogenetic problem-solving. To do that, we are planning to also extend the capabilities of our problem-solving environment. Our most ambitious plan would be to create two or more data sets that include morphological, biogeographical, ontogenetic, stratigraphic, karyotypic, and molecular data. One of these data sets would be based on an imaginary set of organisms, possibly the Caminalcules. Others would be real data sets of well-studied groups. In conjunction with the data sets would be a set of software tools for viewing and comparing data from different sources and using this data to construct phylogenetic histories. We have thought of many ways that an environment like this could be used. For example, students could be divided into groups that has access to fossils and living organisms from a restricted geographic area. Subsequently, from the individual studies the class would be challenged to try to construct a unified phylogenetic history of the group. The possibilities are endless.

We're looking for people who would be willing to collaborate with us on a project like this. We need people who can increase the biological background and content of our project and who can help us find, create, and organize phylogenetic data sources. Most importantly (from my standpoint, anyway) I need experts who would be willing to volunteer to solve a few problems using Phylogenetic Investigator and help me to gather data for my dissertation.

My dissertation examines how experts construct phylogenetic trees from coded and polarized data. I'm studying experts in order to develop a model of desired performance for students -- a model that describes what content knowledge, strategic

knowledge, and organization of knowledge is desirable for students to achieve in this domain. This model is based partly on textbooks and methodological articles, but a careful study of experts greatly increases the trustworthiness by providing evidence of the kinds of mental processes that practitioners actually use when working these problems. If you perform phylogenetic analysis as part of your regular research, I hope you will seriously consider volunteering for my study.

I'm planning to make Phylogenetic Investigator available free of charge (as 'freeware') to anyone who is interested in using it in the classroom. Please reply to me via email if you're interested in finding out more about Phylogenetic Investigator, potentially collaborating with us, or (please oh please) participating in my study.

To find out more and/or volunteer, please call me (616-387-7638) or send me some email (brewer@cs.wmich.edu). We also have a website from which you can download Phylogenetic Investigator, my dissertation proposal, and other supporting documents. Point your web-browser to: <http://141.218.91.93/PIGuide/piguide.html>

Appendix D
Letter of Recruitment

Letter of Recruitment

I am looking for phylogenetic biologists to participate in a study I am conducting for my dissertation. I hope to improve the teaching of systematic and evolutionary biology by having students solve realistic problems of phylogenetic inference. My study, which examines how experts solve these problems, will be useful for developing curricula and instructional materials. If you agree to participate, I will ask you to volunteer about hours of your time to become familiar with the software I am using and to draw a series of phylogenetic trees using coded and polarized data that I will provide.

Although I can't offer any direct compensation, you may keep the software we use for teaching purposes, if you like. In addition, this may be an opportunity for you to think about your teaching in a way you haven't considered before. Finally, if you find our work interesting, we are setting up a long-term project that will use this study as a foundation: We are looking for people who would be willing to collaborate in creating a microworld that students could explore that is composed of multiple data sources. We think that this project has the potential to be a powerful tool for teaching systematics and evolutionary biology.

To find out more and/or volunteer, please call me (616-387-7638) or send me some email (brewer@cs.wmich.edu).

Appendix E
Research Checklist

Research Checklist

Set up Computer

Set up Tape Recorder

Sign Letter of Consent

Demonstrate Program

Read Directions:

I'm going to give you some problems of phylogenetic tree construction. The characters and states of the organisms have already been selected and their polarity has been established. All you have to do is use this data to construct what you believe to be the best phylogenetic tree. My goal in having you do this is to find out what you're thinking while you're solving these problems. While you work on the problems, I'd like you talk as constantly as you can. Try to simply speak the words that are passing through your mind as you solve the problems.

You don't need to say anything special or to clarify your thinking. Just work as you normally would. Some people say that they mumble to themselves while they are solving problems. If that's what you do, then just mumble louder. In any case, try to talk constantly. Say what you are thinking and doing even if it doesn't make sense.

Let's try a practice problem or two before we really get started. Would you work on it now and talk aloud as much as possible.

Start Taping

Start Practice Problems

Start Research Problems

During Interview:

Can you say what you're thinking?
That's very clear
Please tell me what you're thinking
Mmmm.
OK.

Assign problems

Export trees when best tree reached (repeat as necessary)

Appendix F
Letter of Informed Consent

Informed Consent

Western Michigan University
Department of Science Studies
Researcher: Steven D. Brewer
Advisor: Dr. Robert Hafner

I have been invited to participate in a research project entitled "Expert Phylogenetic Tree Construction using Phylogenetic Investigator." I understand that this study seeks to describe what knowledge and organization of knowledge experts use to construct phylogenetic trees. I recognize that this study is Steve Brewer's dissertation project.

I agree to meet the researcher at a location of our mutual agreement for a four to six hour problem-solving session. At this meeting, I will learn how to use *Phylogenetic Investigator* (PI) and the "think-aloud" protocol. I will then solve a series of problem of phylogenetic tree construction using PI and the think-aloud protocol. After each problem, the researcher may ask clarifying questions. After the session has been completed, I will also be asked to evaluate the extent to which these problems caused me to engage in the kinds of thinking I usually use when solving problems of phylogenetic tree construction.

I am aware that while I am solving these problems, PI will be recording the actions I take in the computer environment. I understand that my think-aloud protocol and retrospective interview responses will be tape-recorded.

I understand that no risks, hazards, or discomforts are foreseen as a consequence of this study. As in all research, there may be unforeseen risks to the participant. If an accidental injury occurs, appropriate emergency measures will be taken; however, no compensation or treatment will be made available to the subject except as otherwise stated in this consent form.

I may benefit from participating in this study by learning how to use PI and by seeing how realistic problems of phylogenetic tree construction can be simulated. I am also aware that knowledge gained from this study may contribute to improving instructional materials for teaching evolutionary biology and systematics.

I understand that, although no sensitive information is being recorded, the information collected from me is confidential: my name will not appear on any papers on which this information is recorded. Subjects will be identified with a coded reference and a master list that shows corresponding names of subjects will be kept separately from the data. This master list will be destroyed at such time as the project is completed. I further understand that all of the collected information will be retained in Dr. Robert Hafner's files for a period of three years after completion of the study.

I understand that I may refuse to participate or quit at any time during the study without penalty or prejudice. If I have questions about the study, I may contact either Steve Brewer at (616) 387-5621 or Dr. Robert Hafner at (616) 387-5844. I may also contact the Chair of Human Subjects Institutional Review Board at (616) 387-8293 or the Vice President for Research at (616) 387-8298 with any concerns that I may have. My signature below indicates that I understand the purpose and requirements of the study and that I agree to participate.

Signature

Date

BIBLIOGRAPHY

- Alters, B. J., & McComas, W. F. (1994). Punctuated equilibrium: The missing link in evolution education. American Biology Teacher, 56(6), 334-340.
- Anzai, Y. (1991). Learning and use of representations for physics expertise. In K. A. Ericsson & J. Smith (Eds.), Toward a general theory of expertise: Prospects and limits (pp. 64-93). Cambridge: Cambridge University Press.
- Bédard, J., & Chi, M. T. H. (1992). Expertise. Current directions in psychological science, 1(4), 135-139.
- Bishop, B. A., & Anderson, C. W. (1986). Student conceptions of natural selection and its roles in evolution (Research Series No. 165). Institute for Research on Teaching, Michigan State University.
- Bock, W. J. (1969). Non validity of the "phylogenetic fallacy". Systematic Zoology, 18(1), 111-115.
- Brooks, D. R., & McLennan, D. A. (1991). Phylogeny, Ecology, and Behavior. Chicago: University of Chicago Press.
- Brooks, D. R., McLennan, D. A., Carpenter, J. M., Weller, S. G., & Coddington, J. A. (1995). Systematics, ecology and behavior. Bioscience, 45(10), 687-695.
- Brooks, D. R & Wiley, E. O. (1986). Evolution as Entropy: Toward a unified theory of biology. Chicago: University of Chicago Press.
- Brooks, D. R & Wiley, E. O. (1988). Evolution as Entropy: Toward a unified theory of biology. 2d. Ed. Chicago: University of Chicago Press.
- Calley, J. N., Jungck, J. R., & Collins, A. (1994). Genetics Construction Kit. In J. R. Jungck (Ed.) BioQUEST Library. University of Maryland Press.
- Carson, H. L. (1975). The genetics of speciation at the diploid level. American Naturalist, 109(965), 83-92.
- Carson, H. L. (1985). Unification of speciation theory in plants and animals. Systematic Botany, 10, 380-390.
- Carson, H. L. (1994). Genetic imbalance, realigned selection, and the origin of species. In Genetic changes in natural populations (pp. 345-362).
- Charness, N. (1989). Expertise in chess and bridge. In D. Klahr & K. Kotovsky (Eds.), Complex information processing: The impact of Herbert A. Simon (pp. 183-208). Hillsdale: Lawrence Erlbaum Associates, Inc.

- Chase, W. G. and Simon, H. A. (1973). Perception in chess. Cognitive Psychology, 4, 55-81.
- Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. Cognitive Science, 5, 121-152.
- Clark, M. (1989). Biological and health sciences: Report of the project 2061 phase I biological and health sciences panel. Washington, DC: American Association for the Advancement of Science.
- Colless, D. H. (1967). The phylogenetic fallacy. Systematic Zoology, 16(4), 289-295.
- Colless, D. H. (1969). The phylogenetic fallacy revisited. Systematic Zoology, 18(1), 115-126.
- Collins, A. (1986) Strategic knowledge required for desired performance in solving transmission genetics problems. Doctoral Dissertation, University of Wisconsin-Madison.
- Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L. B. Resnick (Eds.), Knowing, learning, and instruction: Essays in honor of Robert Glaser Hillsdale: Lawrence Erlbaum Associates, Inc.
- Council, N. R. (1990). Fulfilling the promise: Biology education in the nation's schools. Washington, DC: National Academy Press.
- Cowan, D. (February 1995) Personal Communication.
- Darden, L. (1991). Theory change in science: Strategies from Mendelian genetics. New York: Oxford University Press.
- Davis, G. M. (1995). Systematics and public health. Bioscience, 45(10), 705-714.
- de Kleer, J., & Brown, J. S. (1983). Assumptions and ambiguities in mechanistic mental models. In D. Gentner & A. L. Stevens (Eds.), Mental Models (pp. 155-190). Hillsdale: Lawrence Erlbaum Associates, Inc.
- de Queiroz, K. (1985). The ontogenetic method for determining character polarity and its relevance to phylogenetic systematics. Systematic Zoology, 34(3), 280-299.
- Dobzhansky, T. (1973). Nothing in biology makes sense except in the light of evolution. American Biology Teacher, 35, 125-129.
- Duncan, T., Phillips, R. B., & W. H. Wagner, J. (1980). A comparison of branching diagrams derived by various phenetic and cladistic methods. Systematic Botany, 5(3), 264-293.
- Duveen, J., & Solomon, J. (1994). The great evolution trial: Use of role play in the classroom. Journal of Research in Science Teaching, 31(5), 575-582.

- Eldredge, N., & Cracraft, J. (1980). Phylogenetic Patterns and the Evolutionary Process: Method and theory in comparative biology. New York: Columbia University Press.
- Eldredge, N., & Gould, S. J. (1972). Punctuated equilibria: An alternative to phyletic gradualism. In T. J. M. Schopf (Eds.), Models in paleobiology (pp. 82-115). San Francisco: Freeman, Cooper, and Co.
- Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. Psychological Review, 87(3), 215-251.
- Ericsson, K. A., & Simon, H. A. (1993). Protocol Analysis: Verbal reports as data. Cambridge: MIT Press.
- Ericsson, K. A., & Smith, J. (1991). Empirical studies of expertise: Prospects and limits. In K. A. Ericsson & J. Smith (Eds.), Toward a general theory of expertise: Prospects and limits (pp. 1-38). Cambridge: Cambridge University Press.
- Estabrook, G. F. (1968). A general solution in partial orders for the Camin-Sokal model in phylogeny. Journal of Theoretical Biology, 21, 421-438.
- Estabrook, G. F. (1977). Does common equal primitive? Systematic Botany, 2, 36-42.
- Estabrook, G. F. (1978). Some concepts for the estimation of evolutionary relationships in systematic botany. Systematic Botany, 3(2), 146-158.
- Estabrook, G. F. (1986). Evolutionary classification using convex phenetics. Systematic Zoology, 35(4), 560-570.
- Estabrook, G. F., & Anderson, W. R. (1978). An estimate of phylogenetic relationships within the genus *Crusea* (Rubiaceae) using character compatibility analysis. Systematic Botany, 3(2), 179-196.
- Farris, F. J. (1977). Phylogenetic analysis under Dollo's law. Systematic Zoology, 22, 50-54.
- Farris, J. S. (1969). A successive approximations approach to character weighting. Systematic Zoology, 18, 374-385.
- Farris, J. S. (1982). The logical basis of phylogenetic analysis. In N. Platnick & V. Funk (Eds.), Advances in cladistics: Proceedings of the Willi Hennig society (pp. 7-36). Columbia University Press.
- Farris, J. S. (1985). Distance data revisited. Cladistics, 1(1), 67-85.
- Farris, J. S. (1986). Distances and statistics. Cladistics, 2(2), 144-157.
- Felsenstein, J. (1983). Parsimony in systematics: Biological and statistical issues. Annual Review of Evolution and Systematics, 14, 313-333.

- Felsenstein, J. (1986). Distance methods: A reply to Farris. Cladistics, 2(2), 130-143.
- Felsenstein, J. (1993). PHYLIP: Phylogeny inference package. University of Washington. Distributed by the author.
- Fifield, S., & Fall, B. (1992). A hands-on simulation of natural selection in an imaginary organism, *Platysoma apoda*. American Biology Teacher, 54(4), 230-235.
- Fitch, W. M. (1971). Toward defining the course of evolution. Systematic Zoology, 20, 406-416.
- Forey, P. L., Humphries, C. J., Kitching, I. J., Scotland, R. W., Siebert, D. J., & Williams, D. M. (1992). Cladistics: A practical course in systematics. Oxford: Clarendon Press.
- Funk, V. A. (1985). Phylogenetic patterns and hybridization. Annual of the Missouri Botanical Garden, 72, 681-715.
- Garnham, A., & Oakhill, J. (1994). Thinking and Reasoning. Oxford: Blackwell.
- Gick, M. L. and Holyoak, K. J. (1983). Schema induction and analogical transfer. Cognitive Psychology, 15, 1-38.
- Gillhooly, K., Wood, M., Kinnear, P. R., & Green, C. (1988). Skill in map reading and memory for maps. Quarterly Journal of Experimental Psychology, 40A, 87-107.
- Gould, S. J. (1977). Ontogeny and Phylogeny. Cambridge: Belknap Press of Harvard University Press.
- Gould, S. J. (1980). The Panda's Thumb: More reflections in natural history. New York: W.W. Norton and Company.
- Gould, S. J. (1989). Wonderful Life: The Burgess Shale and the nature of history. New York: W.W. Norton and Company.
- Gould, S. J., & Lewontin, R. C. (1978). The spandrels of San Marco and the Panglossian paradigm: A critique of the adaptationist programme. Proceedings of the Royal Society of London, 205, 581-598.
- Greeno, J. G. (1983). Conceptual entities. In D. Gentner & A. L. Stevens (Eds.), Mental Models (pp. 227-252). Hillsdale: Lawrence Erlbaum Associates, Inc.
- Greeno, J. G. and Simon, H. A. (1988). Problem solving and reasoning. In R. C. Atkinson, R. J. Herrnstein, G. Lindzey, & R. D. Luce (Eds.), Stevens' handbook of experimental psychology (Rev. Ed.). New York: Wiley.
- Hafner, R. S. (1991) High school students' model-revising problem solving in genetics. Dissertation, University of Wisconsin-Madison.

- Hafner, R. S. and J. Stewart. (1995) Revising explanatory models to accomodate anomalous genetic phenomena: Problem solving in the "context of discovery". Science Education, 79, 111-146.
- Harper, C. W., Jr. (1976). Phylogenetic inference in paleontology. Journal of Paleontology, 50(1), 180-193.
- Harvey, P. H., & Pagel, M. D. (1991). The Comparative Method in Evolutionary Biology. New York: Oxford University Press.
- Haszprunar, G. (1992). How reliable are computer parsimony programs for systematics? An addition. Zietschrift de zoologische und Systematik Evolution-forschung, 30, 244-248.
- Heller, J., & Reif, F. (1984). Prescribing effective human problem solving processes: Problem description in physics. Cognition and Instruction, 1(2), 177-216.
- Hennig, W. (1966). Phylogenetic Systematics. Chicago: University of Illinois Press.
- Huelsenbeck, J. P. (1991). When are fossils better than extant taxa in phylogenetic analysis. Systematic Zoology, 40(4), 458-469.
- Hull, D. L. (1988). Science as a Process. Chicago: University of Chicago Press.
- Juncgk, J., & Calley, J. (1984). Genetics Construction Kit.
- Jungck, J. & Calley, J. (1985). Strategic simulations and post-socratic pedagogy: Construction computer software to develop long-term inference through experimental inquiry. American Biology Teacher. 14(2), 137-146.
- Kindfield, A. C. H. (1994a). Biology diagrams: Tools to think with. Journal of Learning Science, 3(1), 1-36.
- Kindfield, A. C. H. (1994b). Understanding a basic biological process: Expert and novice models of meiosis. Science Education, 78(3), 255-283.
- Kitcher, P. (1982). Abusing Science: The case against creationism. Cambridge: MIT Press.
- Kluge, A. G., & Farris, J. S. (1969). Quantitative phyletics and the evolution of anurans. Systematic Zoology, 18, 1-32.
- Kotovskiy, K., & Fallside, D. (1989). Representation and transfer in problem solving. In D. Klahr & K. Kotovsky (Eds.), Complex information processing: The impact of Herbert A. Simon (pp. 69-108). Hillsdale: Lawrence Erlbaum Associates, Inc.
- Larkin, J. (1983). The role of problem representation in physics. In D. Gentner & A. L. Stevens (Eds.), Mental models (pp. 75-98). Hillsdale: Lawrence Erlbaum Associates, Inc.

- Larkin, J., & Rainard, B. (1984). A research methodology for studying how people think. Journal of Research in Science Teaching, 21(3), 235-254.
- Larkin, J., & Reif, F. (1979). Understanding and teaching problem solving in physics. European Journal of Science Education, 1(2), 191-203.
- Larkin, J. H. (1989). Display-based problem solving. In D. Klahr & K. Kotovsky (Eds.), Complex information processing: The impact of Herbert A. Simon (pp. 319-341). Hillsdale: Lawrence Erlbaum Associates, Inc.
- Larkin, J. H., & Simon, H. A. (1987). Why a diagram is sometimes worth ten thousand words. Cognitive Science, 11, 65-99.
- Lauder, G. V., Huey, R. B., Monson, R. K., & Jensen, R. J. (1995). Systematics and the study of organismal form and function. Bioscience, 45(10), 696-704.
- Lieberman, B. S., & Vrba, E. S. (1995). Hierarchy theory, selection, and sorting. Bioscience, 45(6), 394-399.
- Lorenzen, S. (1992). PHYLIP, PAUP, and HENNIG-86 need improvement. Zietschrift de zoologische und Systematik Evolution-forschung, 30, 249-255.
- Lorenzen, S. (1993). The role of parsimony, outgroup analysis, and theory of evolution in phylogenetic systematics. Zietschrift de zoologische und Systematik Evolution-forschung, 31, 1-20.
- Lorenzen, S., & Sieg, J. (1991). PHYLIP, PAUP, and HENNIG-86 -- how reliable are computer parsimony programs used in systematics? Zietschrift de zoologische und Systematik Evolution-forschung, 29, 466-472.
- Maddison, W., & Maddison, D. (1989). MacClade: Software for cladistic analysis. Sinauer Associates, Inc.
- Maddison, W. P., Donoghue, M. J., & Maddison, D. R. (1984). Outgroup analysis and parsimony. Systematic Zoology, 33(1), 83-103.
- Mayr, E. (1988). Toward a new philosophy of biology: Observations of an evolutionist. Cambridge: Harvard University Press.
- Mayr, E. (1991). One long argument: Charles Darwin and the genesis of modern evolutionary thought. Cambridge: Harvard University Press.
- Meacham, C. A., & Estabrook, G. F. (1985). Compatibility methods in systematics. Annual Review of Ecological Systematics, 16, 431-446.
- Meier, R., & Whiting, M. F. (1992). Hennig86 and PAUP are reliable: A reply to Lorenzen and Sieg. Zietschrift de zoologische und Systematik Evolution-forschung, 30, 239-243.
- Miller, D. R., & Rossman, A. Y. (1995). Systematics, biodiversity, and agriculture. Bioscience, 45(10), 680-686.

- Minitab (1992). MINITAB: Software for statistical analysis. State College, PA: Minitab, Inc.
- Miyamoto, M. M., & Cracraft, J. (1991). Phylogenetic analysis of DNA sequences. New York: Oxford University Press.
- National Research Council (1990). Report of the Committee on High-School Biology Education. Washington DC: National Academy Press.
- Nelson, G. (1978). Ontogeny, phylogeny, paleontology, and the biogenetic law. Systematic Zoology, 27, 324-345.
- Nelson, G. J. (1971). "Cladism" as a philosophy of classification. Systematic Zoology, 20(3), 373-376.
- Nelson, G. J. (1972a). "Science or politics?" A reply to H.F. Howden. Systematic Zoology, 21(3), 341-342.
- Nelson, G. J. (1972b). Phylogenetic relationship and classification. Systematic Zoology, 21(3), 227-231.
- Nelson, G. J. (1972c). Review of Die Rekonstruktion der Phylogenese mit Hennig's Prinzip. Systematic Zoology, 21(3), 350-352.
- Nelson, G. J. (1973). The higher-level phylogeny of vertebrates. Systematic Zoology, 22(1), 87-91.
- Newell, A. (1980). One final word. In D. Tuma & F. Reif (Eds.), Problem solving and education: Issues in teaching and research Hillsdale: Lawrence Erlbaum Associates, Inc.
- Newell, A. (1990). Unified Theories of Cognition. Cambridge: Harvard University Press.
- Nickles, T. (1981). What is a problem that we may solve it? Synthese, 47, 85-118.
- Norman, D. A. (1983). Some observations on mental models. In D. Gentner & A. L. Stevens (Eds.), Mental models (pp. 7-14). Hillsdale: Lawrence Erlbaum Associates, Inc.
- O'Hara, R. J. (1988). Homage to Clio, or, Toward an historical philosophy for evolutionary biology. Systematic Zoology, 37(2), 142-155.
- O'Hara, R. J. (1991). Representations of the natural system in the Nineteenth Century. Biology and Philosophy, 6, 255-274.
- O'Hara, R. J. (1992). Telling the tree: Narrative representation and the study of evolutionary history. Biology and Philosophy, 7, 135-160.
- O'Hara, R. J. (1993). Systematic generalization, historical fate, and the species problem. Systematic Biology, 42(3), 231-246.

- Offner, S. (1994a). Using chromosomes to teach evolution. II. Chromosomal rearrangements in speciation events. American Biology Teacher, 56(2), 86-93.
- Offner, S. (1994b). Using chromosomes to teach evolution. I. Conserved genes and gene families. American Biology Teacher, 56(2), 79-85.
- Olsen, J. R., & Biolsi, K. J. (1991). Techniques for representing expert knowledge. In K. A. Ericsson & J. Smith (Eds.), Toward a general theory of expertise: Prospects and limits (pp. 240-285). Cambridge: Cambridge University Press.
- Platnick, N. I., Griswold, C. E., & Coddington, J. A. (1991). On missing entries in cladistic analysis. Cladistics, 7, 337-343.
- Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. Science Education, 66(2), 211-227.
- Priest, A. G., & Lindsay, R. O. (1992). New light on novice-expert differences in physics problem solving. British Journal of Psychology, 83, 389-405.
- Reif, F. (1983). Understanding and teaching problem solving in physics. In Lectures at the International Summer School on Physics Education, La Londe Les Maures, France:
- Ridley, M. (1986). Evolution and Classification: The reformation of cladism. New York: Longman Group Limited.
- Rutherford, F. J., & Ahlgren, A. (1990). Science for All Americans. New York: Oxford University Press.
- Rutowski, R. L. (1992). Two issues in evolution education. In R. J. Good, J. E. Trowbridge, S. S. Demastes, J. H. Wandersee, M. S. Hafner, & C. L. Cummins (Eds.), Evolution Education Research Conference, (pp. 166-169). Louisiana State University:
- Sanderson, M. J., Baldwin, B. G., Bharathan, G., Cambell, C. S., Dohlen, C. V., Ferguson, D., Porter, J. M., Wojciechowski, M. F., & Donoghue, M. J. (1993). The growth of phylogenetic information and the need for a phylogenetic data base. Systematic Biology, 42(4), 562-568.
- Savage, J. M. (1995). Systematics and the biodiversity crisis. Bioscience, 45(10), 673-679.
- Scott-Ram, N. R. (1990). Transformed cladistics, taxonomy, and evolution. Cambridge: Cambridge University Press.
- Sharmann, L. G. (1993). Teaching evolution: Designing successful instruction. American Biology Teacher, 55(8), 481-486.
- Simpson, B. B., & Cracraft, J. (1995). Systematics: The science of biodiversity. Bioscience, 45(10), 670-672.

- Skoog, G. (1992). Problems associated with evolution education. In R. J. Good, J. E. Trowbridge, S. S. Demastes, J. H. Wandersee, M. S. Hafner, & C. L. Cummins (Ed.), Evolution Education Research Conference, (pp. 181-183). Louisiana State University:
- Sober, E. (1988). Reconstructing the Past: Parsimony, evolution and inference. Cambridge: MIT Press.
- Sober, E. (Ed.). (1994). Conceptual Issues in Evolutionary Biology. Cambridge: MIT Press.
- Sokal, R. R. (1983a). A phylogenetic analysis of the Caminalcules: I. The data base. Systematic Zoology, 32(2), 159-184.
- Sokal, R. R. (1983b). A phylogenetic analysis of the Caminalcules: II. Estimating the true cladogram. Systematic Zoology, 32(2), 185-201.
- Soloway, E., & Ehrlich, K. (1984). Empirical studies of programming knowledge. IEEE Transactions on Software Engineering, SE-10(5), 595-609.
- Stevens, P. F. (1991). Character states, morphological variation, and phylogenetic analysis: A review. Systematic Botany, 16, 553-583.
- Stewart, J. (1988). Potential learning outcomes from solving genetics problems: A typology of problems. Science Education, 72(2), 237-254.
- Stewart, J., & Hafner, R. S. (1994). Research on problem solving: Genetics. In D. L. Gable (Eds.), Handbook of research on science teaching and learning (pp. 284-200). New York: MacMillan Publishing Co.
- Stewart, J., & Jungck, J. R. (1994). Problem-posing, Problem-solving, and Persuasion in Biological Investigations. In J. R. Jungck (Ed.) BioQUEST Library. University of Maryland Press.
- Stuessy, T. F., & Crisi, J. V. (1984). Problems in the determination of evolutionary directionality of character-state change for phylogenetic reconstruction. In T. Duncan & T. F. Stuessy (Eds.), Cladistics: Perspectives on the reconstruction of evolutionary history (pp. 71-87). New York: Columbia University Press.
- Swofford, D. L. (1991). PAUP: Phylogenetic Analysis Using Parsimony. Champaign: Illinois Natural History Survey.
- Thompson, P. (1989). The structure of biological theories. Albany: State University of New York Press.
- Trowbridge, J. E., & Wandersee, J. H. (1994). Identifying critical junctures in a college course on evolution. Journal of Research in Science Teaching, 31(5), 459-473.

- van Someren, M. W., Barnard, Y. F., & Sandberg, J. A. C. (1995). The Think Aloud Method: A practical guide to modelling cognitive processes. New York: Academic Press.
- Vogt, K. D. (1995). Demonstrating biological classification using a simulation of natural taxa. American Biology Teacher, 57(5), 282-283.
- Voss, J. F. (1989). Problem solving and the educational process. In A. Lesgold & R. Glaser (Eds.), Foundations for a psychology of education (pp. 251-294). Hillsdale: Lawrence Erlbaum Associates, Inc.
- Wägele, J. W. (1994). Review of methodological problems of "Computer cladistics" exemplified with a case study on isopod phylogeny (Crustacea:Isopoda). Zeitschrift für zoologische Systematik und Evolutionsforschung, 32, 81-107.
- Wagner, W. H. (1980). Origin and philosophy of the groundplan-divergence method of cladistics. Systematic Botany, 5(2), 173-193.
- Wagner, W. H. (1983). Reticulistics: The recognition of hybrids and their role in cladistics and classification. In N. I. Platnick & V. A. Funk (Ed.), Second meeting of the Willi Hennig society, . Columbia University Press.
- Wagner, W. H., & Beitel, J. M. (1992). Generic classification of modern North American Lycopodiaceae. Annals of the Missouri Botanical Garden, 79, 676-686.
- Waterman, M. S. (1995). Introduction to computational biology: Maps, sequences and genomes. New York: Chapman & Hall.
- Wiley, E. O. (1980). Phylogenetic systematics and vicariance biogeography. Systematic Botany, 5(2), 194-220.
- Wiley, E. O. (1981). Phylogenetics: The principles and practice of phylogenetic systematics. New York: John Wiley & Sons.
- Wiley, E. O., Siegal-Causey, D., Brooks, D. R., & Funk, V. A. (1991). The Compleat Cladist (Special Publication No. 19). University of Kansas Museum of Natural History.
- Williams, M. D., Hollan, J. D., & Stevens, A. L. (1983). Human reasoning about a simple physical system. In D. Gentner & A. L. Stevens (Eds.), Mental Models (pp. 131-154). Hillsdale: Lawrence Erlbaum, Inc.
- Wimsatt, W. C. (1986). Heuristics and the study of human behavior. In D. W. Fiske & R. A. Sweder (Eds.), Methodology in the social sciences: Pluralisms and subjectivities (pp. 293-314). Chicago: University of Chicago Press.
- Wimsatt, W. C. (1987). False models as means to truer theories. In M. Nitecki (Eds.), Neutral models in biology (pp. 23-55). London: Oxford University Press.