TAXONOMY OF COMPUTER SCIENCE & ENGINEERING

Compiled by the
AFIPS Taxonomy Committee

TAXONOMY OF COMPUTER SCIENCE & ENGINEERING

Compiled by the
AFIPS Taxonomy Committee

Robert L. Ashenhurst
Harold Borko
Sam D. Conte
Patrick Fischer
Bernard A. Galler
George Glaser
Anthony Ralston
(Chairman)
Edwin D. Reilly
Gerard Salton
Jean E. Sammet
Stephen S. Yau

with the help of
over 70 consultants

Printed in the United States of America

Library of Congress Catalog Card No. 79-57474

ACKNOWLEDGEMENTS

This project has been funded by:

Institute for Computer Sciences and Technology of the National Bureau of Standards

Air Force Office of Scientific Research

Office of Naval Research

AFIPS is most appreciative to these agencies for their support.

Copious thanks are also due to Carol Harter and Janet Root who typed the many iterations of this report and maintained the file from which the Taxonomy Tree and Annotated Taxonomy Tree were produced camera-ready for this book.

PREFACE

This document is the result of three years work by a committee of eleven with the aid of over seventy others who either authored part of the *Taxonomy* or reviewed the work of the authors. The result is a study of the structure of a discipline which appears to be unique among the sciences and almost unique among all disciplines.

Why should computer scientists and engineers, whose discipline is the newest among all the scientific and technical disciplines, embark on such a project when their confreres in older, more established disciplines have not done so? The essential reason is lack of understanding of what computer science and engineering is by those outside the discipline. The newness of the field is one reason for this; the other is its very rapid growth, faster, in fact, by any measure than the growth at any time of any other discipline in the sciences or humanities. And, if the pace at the beginning of the fourth decade of the existence of computers is not as frantic as in the first three, it is still so rapid that even those within the discipline are hard pressed to be conversant with anything but a narrow specialty.

But does it matter—except perhaps to our self-esteem—that we are ill-understood or even misunderstood by educated laymen, by government and even by other scientists and engineers? We think it does. Our impact in late twentieth century society is already large and will become immense. If universities still do not always understand the need to provide (informed) instruction in our subject, if governments do not understand our role in their own efficient functioning or in the economics they oversee—to mention only two aspects of the problem—the results can be serious—or worse. This *Taxonomy* will not solve these problems. But as it contributes to increased understanding and heightened perception of what we are and what we do, it will have served a useful role.

For any discipline, but particularly for one changing so rapidly, a taxonomy is only a snapshot, a picture at a particular time. That time for this document is (early) 1979 although, to be fair, almost all of it was produced in 1977 and 1978. If this document proves to be a useful one, then it will need periodic revision. Moreover, the first attempt at a task of this magnitude is surely imperfect in a variety of ways. We would, therefore, like to encourage all readers to communicate to us

—any errors or inaccuracies which you may find (or suspect)
—omissions which you believe are significant
—comments aimed at improving the presentation or annotations.

These should be sent to:

Chairman, Taxonomy Committee
American Federation of Information Processing Societies
1815 North Lynn Street
Arlington, VA 22209

8360108

LIST OF CONTRIBUTORS

Taxonomy Committee

Robert L. Ashenhurst
University of Chicago

Harold Borko
UCLA

Sam D. Conte
Purdue University

Patrick Fischer
Vanderbilt University

Bernard A. Galler
University of Michigan

George Glaser
Centigram Corporation

Anthony Ralston (Chairman)
SUNY at Buffalo

Edwin D. Reilly
SUNY at Albany

Gerard Salton
Cornell University

Jean E. Sammet
IBM Corporation

Stephen S. Yau
Northwestern University

Other Authors

Jon T. Butler
Northwestern University

Billy Claybrook
University of South Carolina

Gerald Engel
Christopher Newport College

Carl Engelman
Mitre Corporation

T. Y. Feng
Wright State University

Aaron Finerman
University of Michigan

Fred H. Harris
University of Chicago

Harold Highland
SUNY at Farmingdale

Laveen Kanal
University of Maryland

Rob Kling
University of California at Irvine

Franklin F. Kuo
University of Hawaii

C. L. Liu
University of Illinois, Champaign

Harold Lorin
IBM Corporation

Joel Moses
M.I.T.

Susan Nycum
Attorney, Chickering & Gregory

Vaughn Pratt
M.I.T.

C. V. Ramamoorthy
University of California at Berkeley

Fred T. Riggs
University of Hawaii

Saul Rosen
Purdue University

Arnold Rosenberg
IBM Corporation

Azriel Rosenfeld
University of Maryland

Paul Roth
U.S. Department of Energy

Stuart Shapiro
SUNY at Buffalo

Ben Shneiderman
University of Maryland

Adrian Stokes
Hatfield Polytechnic, England

Jean-Paul Tremblay
University of Saskatchewan

Andries Van Dam
Brown University

Karl Zinn
University of Michigan

Reviewers and Consultants

John Berg
National Bureau of Standards

Robert Bigelow
Attorney, Bigelow & Saltzberg

Abraham Bookstein
University of Chicago

Frederick M. Brasch
Northwestern University

Fred Brooks
University of North Carolina

Thomas Cain
University of Pittsburgh

Richard Canning
Canning Publications

John Case
SUNY at Buffalo

Steven Coons
University of Colorado, Boulder

David Cooper
University College, London

Fernando Corbató
M.I.T.

Ira Cotton
National Bureau of Standards

Donald Crouch
University of Alabama

Philip Dorn
Dorn Computer Consultants

Marvin Ehlers
Square D Corporation

Donald Frazer
IBM Corporation

John Goodenough
Softech, Inc.

C. C. Gotlieb
University of Toronto

James Griesmer
IBM Corporation

Michael Hammer
M.I.T.

David Hsaio
Ohio State University

Edwin Istvan
National Bureau of Standards

Richard Karhuse
Northwestern University

John Killeen
University of California at Davis

Peter Kirstein
University College, London

Daniel Klassen
Minnesota Educational Computing Consortium

Jon Meads
Intel Corporation

Cleve Moler
University of New Mexico

Abbe Mowshowitz
University of British Columbia

Bary Pollack
University of British Columbia

Udo Pooch
Texas A&M University

Jeffrey Posdamer
SUNY at Buffalo

Keith Price
University of Southern California

Robert Rosin
Bell Laboratories

Sartaj Sahni
University of Minnesota

Robert Sargent
Syracuse University

Alex Shistakov
Lawrence Livermore Laboratory

Susan Solomon
Eastern Washington University

Sargur Srihari
SUNY at Buffalo

Richard Stark
University of New Mexico

Henry Tropp
Humboldt State University, California

Stephen Unger
Columbia University

Terry Winograd
Stanford University

Patrick Winston
M.I.T.

TABLE OF CONTENTS

I. INTRODUCTION AND GUIDE

If the purpose of producing a taxonomy of computer science and engineering (CS&E) were to be summarized in one word, that word would be *perspective*. The rate of change and progress in CS&E has been so rapid that it has been very difficult to obtain a coherent view of what we are and what we do from outside the discipline and even, sometimes, from within it. If, then, this document, snapshot though it is, can provide a wide-angle view of CS&E, it may help to achieve the perspective of our discipline which is necessary if society is going to use well and wisely the products of CS&E and if the practitioners of the discipline itself are to be useful advisers on applications of and new directions for computer technology.

The starting point in the development of a taxonomy of any subject must be a definition of what that subject encompasses. It seems to us, perhaps parochially, that the boundaries of CS&E are more difficult to ascertain than those of more established disciplines. Partly this results from the rapid growth alluded to above; partly it is because our discipline impinges on areas of more established ones in ways which make it difficult to determine where many topics should be classified. We do not wish to appear to be empire builders gathering into the bosom of CS&E ever larger areas of knowledge and activity. Neither do we wish to appear to be so timid that we do not include areas where, whatever the history, the main contributions are now being made by computer scientists and engineers. The definition which follows is, we believe, a reasonable compromise between these extremes.

Definition of Computer Science and Engineering

Computer science and engineering includes all subject matter normally subsumed under the following rubrics:

COMPUTER SOFTWARE
COMPUTER HARDWARE
COMPUTER SYSTEMS
THEORY OF COMPUTATION
THE COMPUTER INDUSTRY
HISTORY OF COMPUTERS
EDUCATION IN COMPUTER SCIENCE AND ENGINEERING
LEGAL, MANAGEMENT, PROFESSIONAL AND SOCIETAL ASPECTS OF COMPUTING

and, as indicated, portions of the subject matter of

MATHEMATICS OF COMPUTING
(namely: much but not all of Numerical Mathematics; some of Automata and Switching Theory; only Random Number Generation and related topics from the general domain of Statistics)

COMPUTER METHODOLOGIES
(namely: those methods and techniques with wide applicability, specifically Algebraic Manipulation, Artificial Intelligence, Computer Graphics, Database Management Systems, Image Processing, Information Storage and Retrieval, Pattern Recognition, Simulation, Sorting and Searching)

APPLICATIONS/TECHNIQUES

(those traditionally considered part of the discipline which would disappear or be very severely impeded if the computer disappeared; illustrative of these are Business Data Processing, Scientific and Engineering Data Processing, Computer-assisted Instruction, Text (Word) Processing).

Some discussion of this definition may be useful. First, one might not expect to find topics such as the History of Computers or the Legal Aspects of Computing included in a document which attempts to define the structure of a discipline. We took the point of view, however, that an understanding of what CS&E is requires a knowledge of all areas in which computer scientists and engineers work and publish and that, indeed, these topics and other similar ones in the definition above provide a necessary breadth of view about CS&E.

A second decision was to include very little explicitly about the applications of computers to the solution of the myriad problems to which they are now applied. To have been exhaustive here would have meant that the applications part of the *Taxonomy* would have dwarfed the rest of it. Moreover, this portion would have become dated and inaccurate much more rapidly than we hope will be true of the rest of the document. We have, therefore, with some misgivings, because of the inevitably biased point of view which results, restricted our applications section to a few illustrative areas among those where, without computers, the applications would become immensely difficult, if not impossible.

The *Taxonomy*, which is the heart of this document, may be viewed as an elaboration of this definition. It is intended to enable the educated layman to derive some understanding of what we do and, at the same time, to permit those of us in various corners of the discipline to understand better what others in the discipline are concerned with.

Structure of the Report

The remainder of this document consists of the following major sections (the numbers correspond to the Table of Contents in which this section is I.).

II. An outline of the taxonomy tree. This tree contains up to six levels of depth; in this outline we present only the first two levels for the purpose of providing a broad overview.

III. The complete taxonomy tree.

IV. The complete tree again but this time with annotations of terms interleaved with the tree itself. Our aim has been to provide brief definitions or be otherwise descriptive for all terms whose meaning would not be transparent to the educated layman.

V. A list of core terminology in CS&E with definitions. This includes terms in wide use but which, for one reason or another, do not appear explicitly in the tree.

VI. A list of "umbrella" words which, although widely used, have such a broad connotation that they, too, do not appear explicitly in the tree.

VII. An essay which discusses some philosophical and technical issues which we found difficult or contentious.

VIII. A bibliography, keyed to first or second level nodes, containing references—mainly books—to which the user of this report may go for further information on the subject matter of CS&E.

IX. A list of abbreviations and acronyms used in the *Taxonomy*.

X. An alphabetically ordered index of terms which will enable the user of this taxonomy to find the location of terms in the tree or in the lists described in V and VI above.

Structure of the Taxonomy Tree

The *Taxonomy* presented here is in standard indented outline form. In the parlance of CS&E, our structure is a tree with CS&E the label for the root of the tree and Hardware, Computer Systems, Data, Software, Mathematics of Computing, Theory of Computation, Methodologies, Applications/Techniques, and Computing Milieux being the labels for the "nodes" at the ends of the branches from the root. Each level of indentation of the outline corresponds to nodes joined by branches to the next higher level node. The furthest depth of indentation—4 levels in some places, 5 or 6 in others—corresponds to the leaves of the tree. Of course, despite the fact that nominally we have a tree, CS&E can not, any more than any other discipline, be described in such a structurally clean way. In fact, our tree is not really a tree at all but, because of the cross-references (*see* and *see also*), it is really a graph. That is, the structure consists of nodes and edges between pairs of nodes and contains, as a tree may not, cycles (i.e. paths beginning and ending at the same node).

Which is, of course, what would be expected. The interrelationships between Hardware, Software, Data and Computer Systems, between these topics and the Theory of Computation, between Methodologies and Applications etc., assure that only with ample cross-referencing can the structure be wholly perceived.

But we have no doubts that a tree is the appropriate mechanism to use in displaying the structure. The familiarity of the outline form together with the fact that cross-references are the exception and not the rule means that a tree provides more information in a convenient form than any other mode of presentation which we might have chosen.

A few more notes are in order about the way the tree has been constructed:

1. Some of the higher level nodes are not terms in CS&E itself but rather *covering* nodes whose only purpose is to collect lower level concepts under a single heading (e.g., 2.1 Structure-based Systems; 4.2.3 Human involvement).

2. No node has a single offspring (i.e. branch) emanating from it.

3. The order of presentation of the subnodes of a given node generally has some rationale such as:

—from practical to theoretical (or vice versa)
—alphabetical order.

4. A limit of 6 levels of the tree has been rigorously enforced. This was done to keep the total size of the project under control and because no case was presented to the committee of a need for more levels where the result would not have been a *Taxonomy* too detailed to serve our intended purposes.

5. Nodes labeled "Other" have been used rather freely at levels 2, 3, and 4 to indicate that the other subnodes of the parent node are not an exhaustive subdivision of the label of that parent node. However, it is intended that no major concepts be subsumed under "Other."

Guide to the Taxonomy

Users of this *Taxonomy* will come to it with varying orientations and needs. We list here a few of the more obvious ways in which it may be convenient to make use of the *Taxonomy*.

1. Although this document is not, and is not intended to be used as a dictionary or encyclopedia, it will often be useful to look up a term in the Index and then find its place in the annotated tree for a definition of it or information about it and an indication of where it fits into the larger structure.

2. Where the main interest is in obtaining a broad overview of a particular subdisciplinary area, it will often be best to start with the outline of the tree and then proceed to the unannotated tree. (Note that both the unannotated and annotated trees have running heads corresponding to the first or second level nodes in the outline.)

3. The annotated tree will have its major use for those who wish a view in depth of some (relatively narrow) portion of CS&E. As such, it may provide a starting point for a particular study of some area with the bibliography providing pointers to sources for deeper study.

Other Uses of the Taxonomy

In addition to the direct and rather specific uses of this document described in the Guide above, there are a number of more general uses to which the *Taxonomy* may be put. One such might be in the allocation of grant funds to support basic and/or applied research in computer science and engineering. From the structure presented here, it should be possible to understand better the component subareas of various areas of research. The distinctions made, for example, between various kinds of computer architecture (node 1.3) might be helpful in understanding an author's emphasis on research on one particular kind of architecture. In this context, it is worth noting that, if this *Taxonomy* had been generated 10 years ago, the entire area of computer architecture might have been represented in a different form, leading to a different understanding of its role within computer science and engineering. The implication of this is, of course, that a taxonomy like this one must be a living document, changing and adapting as the discipline develops and changes.

Another use which might be made of this *Taxonomy* is in the classification of jobs within a large organization or within governmental civil service. There are many ways to classify positions in industry or government, but usually there is an implicit underlying classification or taxonomy of the field itself. It may well be the case that the difficulty that the federal government and other governmental bodies and large organizations have had in classifying computer science and engineering positions may arise in large measure from a lack of any clear perception of the discipline of CS&E. We hope this *Taxonomy* will be able to provide such perception. Moreover, the structure and relationship between subfields shown here might very well suggest distinctions and/or similarities between otherwise unrelated (or too closely-related) jobs.

Still another application area for this *Taxonomy* may be in the organization of catalogues, review documents, and journals. *Computing Reviews,* for example, is the leading review journal for computer science and engineering publications; it has its own "review categories," developed some years ago. It may be expected that this *Taxonomy* will have some effect on the evolution of the review categories used by *Computing Reviews* and other, related journals.

Finally, one may expect a journalist or other newcomer to the field to find this *Taxonomy* helpful in obtaining an overview and in learning some of the basic terminology. The kinds of distinctions made in creating a taxonomy are precisely those needed to understand arguments and predictions about past history and future growth and emphasis within a discipline.

A document such as this one has, unfortunately, not only its uses but also its potential abuses. The most likely abuse, perhaps, is that, when it conforms to the notions of those with axes to be ground, it will be quoted as revealed truth. It is, of course, nothing of the kind. Even though there has been widespread input from the CS&E community, this *Taxonomy* is, as we indicate elsewhere, inevitably the result of much compromise. Moreover, as the first attempt at the task of classifying the subject matter of CS&E, it must be considered only a first approximation.

It is also possible to abuse our intent by reading more into this *Taxonomy* than was intended. Readers may tend to measure the relative importance of the topics in the *Taxonomy* by such inappropriate means as the space devoted to each first or second level node. Or they may forget that the rapidly changing structure of CS&E requires all use to take into account the creation date of this document. In general, abuse is likely whenever it is forgotten that this *Taxonomy* is and can be nothing more than a snapshot of the discipline taken from one of several possible angles.

Our hope is that we have created a document which will stimulate readers to find unexpected uses but not abuses of this *Taxonomy*. As these new uses are found and become understood, we shall undoubtedly find ways to improve the structure and the presentation of the information herein. We are looking forward to that task.

A final note. Development of this *Taxonomy* has been a humbling task for all of those who have worked on it. There is, inevitably, no single *right way* to classify the subject matter of CS&E. Within the Committee which oversaw this project, between Committee members and authors of portions of the *Taxonomy*, and between authors and reviewers there have been disputes, sometimes heated ones. Not only is it inevitable that our decisions will not satisfy everyone or, perhaps, anyone; it is also inevitable that there will be omissions, possibly some egregious ones, in this document. To a degree, tasks like this one must be undertaken a first time so that they may be got right the next time. We urgently invite constructive criticism from our readers which will inform the next version of this document.

II. TAXONOMY TREE OUTLINE

1. Hardware

- 1.1 Types of computers
- 1.2 Digital computer subsystems
- 1.3 Digital computer architecture
- 1.4 Input/Output devices
- 1.5 Computer circuitry
- 1.6 Computer elements
- 1.7 Computer hardware reliability

2. Computer Systems

- 2.1 Structure-based systems
- 2.2 Access-based systems
- 2.3 Special purpose systems
- 2.4 Performance of systems

3. Data

- 3.1 Data structures
- 3.2 Data storage representation
- 3.3 Data management
- 3.4 Data communications

4. Software

- 4.1 Tools and techniques
- 4.2 Programming systems
- 4.3 Data and file organization and management

5. Mathematics of Computing

- 5.1 Continuous mathematics
- 5.2 Discrete mathematics
- 5.3 Numerical software and algorithm analysis

6. Theory of Computation

- 6.1 Switching and automata theory
- 6.2 Formal languages
- 6.3 Analysis of programs
- 6.4 Computer models
- 6.5 Complexity of computations
- 6.6 Analysis of algorithms

7. Methodologies

- 7.1 Algebraic manipulation
- 7.2 Artificial intelligence
- 7.3 Information storage and retrieval
- 7.4 Database management systems
- 7.5 Image processing
- 7.6 Pattern recognition
- 7.7 Modeling and simulation
- 7.8 Sorting and searching
- 7.9 Computer graphics

8. Applications/Techniques (Illustrative)

- 8.1 Business data processing
- 8.2 Scientific and engineering data processing techniques
- 8.3 Computer-assisted instruction
- 8.4 Text (word) processing

9. Computing Milieux

- 9.1 The computer industry
- 9.2 Education and computing
- 9.3 History of computing
- 9.4 Legal aspects of computing
- 9.5 Management of computing
- 9.6 The computing profession
- 9.7 Social issues and impacts of computing

III. TAXONOMY TREE