Foundations of logico-linguistics: a unified theory of information, language and logic

W S Cooper

Reidel, USA (1978) 249 pp

In this book, Cooper seeks to provide a framework for a unified science of logic and linguistics. The fundamental primitive concept which he uses to achieve this unification is that of information, and thus his book can also be seen as presenting a theory of information, although since information is the primitive concept, it cannot be defined precisely, but only given a loose intuitive characterization.

Cooper's work is related in approach to two other recent endeavours. As a theory of information and of meaning it is a continuation of the work of Carnap and Bar-Hillel, who sought to provide a semantic theory of information as distinguished from a theory of signal transmission, Bar-Hillel's term for the communications theory of information of Shannon. As Bar-Hillel points out, there is no logical connection between the amount of semantic information in a message and the measure of rarity of the symbol sequences used to convey the message. As a theory of logic, Cooper's work is comparable in intent to that of Russell and Whitehead in *Principia mathematica,* except that where they tried to show that mathematics was a branch of logic, he tries to show that logic is a branch of linguistics. Finally, as a theory of linguistics, Cooper's work is an ambitious attempt to formalize language study in the areas of semantics and pragmatics in a similar fashion to what is now being done in the area of syntax.

Cooper motivates the discussion of his theory by showing how the application of formal logic to ordinary natural language inference leads to obviously wrong results. Mathematical logic, of course, need have no connection with natural language use of logic, but if one wants a comprehensive theory of logic that can, among other things, describe how logic is used in natural language, then one must consider such inconsistencies. Rather than trying to examine the close connection between language and logic noted by many thinkers over the ages, Cooper seeks to develop a unified theory of logico-linguistics.

Information

The theory starts with the concept of information which is characterized by:

- being stored in a physical system,
- that system's having a set of states that it can be in,
- the limitation that the system can only be in one state at a given time, except possibly briefly while in state transition,
- a set of possible information inputs to the system,
- a set of possible test, or question inputs with corresponding outputs.

Given this characterization of information, it is easy to see that this system can be viewed as an automaton. Cooper uses the Mealy model to define this system. Briefly the definition is that there is a 5-tuple $(X, Y, Z, d, 1)$ where X, Y and Z are the nonempty input, output and state sets respectively, and d is a possibly partial state transition function from $Z \times X$ into Z, and 1 a possibly partial output function from $Z \times X$ into Y. The functions may be partial because some functions are undefined; for example, in the case of a language automaton which will shortly be defined, the state in which p and not-p is affirmed would not be defined. An information automaton is an automaton with the additional restriction that the input set X is further divided into X_i, or informative input which has null output, and X_t, or question input which does not cause a state change. Finally, a language automaton is an information automaton with input restricted to linguistic input.

The notion of a language automaton is central to the theory of this book. It can be defined more simply, but equivalently, as a 4-tuple (S, Z, L, B) where S is the sentence set, Z the state set, L the learning operation and B the belief operation. This simplification is possible because S can cover input and output sets, while the possibly partial functions can be restated in terms of learning and belief. If an informative input is accepted, it is learned. If a test input yields a given output, that output is believed in the current state of the system.

Behaviourist definition

Thinking about language in terms of automata theory leads Cooper to take a behaviourist approach to language study. An automaton is a black box. We can observe the input and resulting output, but can only hypothesize about the inner workings. For this reason he introduces the notion of equivalence classes of language automata. Any two automata whose input/output behaviour is indistinguishable are equivalent. If we consider the class of all language automata which are equivalent, this is the definition of a language.

Speakers of a given language can all be thought of as carrying language automata around in their heads which are in an equivalence class for that language. This equivalence explains our ability to communicate with other speakers of our language. Language automata do

not provide the whole explanation of language behaviour, however. A language automaton is only concerned with modelling linguistic competence, not performance, to make the Chomskian distinction which Cooper follows. Furthermore, if we make the traditional breakdown of language study into syntax, semantics and pragmatics, Cooper's theory only extends partially into the realm of pragmatics.

If we try to describe a language in terms of an equivalance class of language automata, how do we know if our description is valid? Basic evidence statements are the main tool for this purpose. We can know that there is an information state z in which S_1 is believed in a given language if for any, and thus every, automaton in the equivalence class representing the language, the automaton, when given some test input, outputs S_1. Here again, as with competence, Cooper follows the current standards of linguistic research and suggests using linguistic informants to carry out such research. One asks informants whether a particular input/output combination could occur.

Inside the black box

Cooper next asks what might go on inside the black box represented by the languae automaton. He introduces the notion of possible worlds as follows. Consider the set of all possible universe histories. Possibility in this case means logical possibility. Thus there could be a world in which things fall upwards, but not one in which all things fall upwards and some fall down. Applying this to the present theory it can be seen that a given fund of information rules out some possible worlds. If one figuratively marks each possible world as ruled out or not, then the worlds and their markings are a representation of that fund of information. Cooper points out that as it stands his use of possible worlds is not too helpful, because the mathematical formalism to represent possible worlds would have to be as complex as possible worlds, in order to represent each possible world uniquely. To get around this problem, one considers two possible worlds linguistically indistinguishable if no sentence holds true in one, but not the other.

Logic

Now that language has been described in terms of language automata, it still remains to be shown how logic fits into this scheme. When a language automaton is given a test input, or question, and yields an answer, it will have to infer an answer which is not explicitly represented in its store of information in its current state. This is where logic enters. Once again the competence distinction is crucial. Unlike a human, a language automaton can carry out logical operations perfectly. The definition of logical implication is as follows: S_1 logically implies S_2 in a language if and only if for each state in the language automaton if S_1 is believed, then so is S_2. Similar definitions can be given for other logical relationships. The deductive system for a given language can be thought of as that ' language's system of implication and logical inconsistency. As before, the accuracy of a language's deductive system is tested by the use of basic evidence statements.

Besides the deductive logic provided by these definitions, inductive logic can be brought in by allowing credibility weights to be used. A language automaton has an input selector which assesses input. It can either accept or reject. Without considering weights, all accepted input is held to be believed and true, but if weights are allowed, it can be believed with a certain weight, or credibility, and if these weights are limited to being between 0 and 1 inclusively, then one can think of them as probabilities. Only logically valid sentences are allowed to be given a weight of 1. Thus the definition of logical implication becomes S_1 logically implies S_2 if and only if, as S_1 approaches 1, S_2 approaches 1, where 1 is the limit towards which a nonlogically valid sentence can tend. Introducing weights also allows a language automaton to come to disbelieve what it formerly believed. Without weights, once a sentence was accepted, it was always believed, but now that a sentence can be accepted at varying credibility levels, further evidence can cause its credibility to diminish to the point where it is no longer believed.

Once one considers the deductive and inductive logical systems of a language, one has fairly completely described the behaviour of the equivalence class of language automata which represent that language. It is for this reason that Cooper asserts his logicolinguistic thesis that 'linguistics (beyond syntax but exclusive of higher pragmatics) and logic are concerned with essentially the same subject matter'.

As an extended example of how logicolinguistic theory can be applied, Cooper compares what he calls hypotheses A and B about the use of 'if . . . then'. Hypothesis A is that the English language use of 'if . . . then' reflects the standard mathematical logic interpretation of material implication. Hypothesis B is that the English language 'if . . . then' behaviour can be understood according to Cooper's logicolinguistic theory. He gives many examples where hypothesis A is found unsuccessful, while hypothesis B is invariably successful.

Computer science

Finally, Cooper considers applications of this theory to computer science and various problems in philosophy and linguistics. The most relevant of these reflections for information retrieval are those concerning computers. A computer is a finite state automaton, and, in fact, is a growing automaton, since an indefinite amount of further memory

or disc storage can be added. Thus, in principle, a computer could be a language automaton. Machine translation, then, to take one example, can be viewed as getting a computer to represent a language automaton which would reflect the competence of a bilingual speaker. Similarly, natural language understanding should be possible; at least in theory.

In conclusion, Cooper's work describes with a high degree of mathematical rigour the foundations for a unified theory of information, language and logic. As these are key concepts in the theory of information retrieval, this book merits attention by those with a theoretical interest in information retrieval. Much more work remains to be done in more fully developing the ideas discussed in this book, but it does represent a carefully thought out starting point.

P Thompson
School of Library and Information Studies, University of California

Automation technology for management and productivity — advancements through CAD/CAM and engineering data handling

P Cheng-Chao Wang (Ed.)

Proceedings of the 3rd Symposium on Automation Technology in Data Handling, Prentice-Hall, USA (1983) £21.20

The reviewer was faced with a confusing task in approaching this book. The publication of the proceedings of a conference should either be an exact copy of what was presented, in the order in which papers were read, thus reflecting the logic of the program organisers, or, preferably, the opportunity should be availed of to reorganize the material into some coherent form, such as is attempted in the publication of any other book. In the case of conference proceedings, this is particularly important to help the reader assimilate the varying styles and approaches of the various authors to a particular topic, and, by their very propinquity, to facilitate the maximization of knowledge from a set of related articles.

This publication fails to achieve these goals for many reasons. The original symposium order of presentation is not retained, as a comparison of the table of contents with the symposium program, shown at the rear of the book, clearly shows. The articles are presented in alphabetical order of principal author name, which consequently randomizes the order of topic presentation. Owing to the author alphabetization order, the keynote address appears at position 15 in the table of contents (and is not indicated as such in this table). In a total of 37 articles, one of which shows no indication of having been presented in any form at the symposium, no less than 11 articles titles have been changed significantly from their original form.

To attempt to remedy the above deficiencies, the reviewer has categorized the articles as follows.

Of the 37 articles, 18 deal with the handling of engineering data, (graphical data, documentation systems, management information systems, database systems etc.), eight deal with graphical techniques (plotting, graphic design, three-dimensional modelling or display etc.), two deal with education (graphics, CAD/CAM etc.), three deal with facility management (some of the 18 articles on engineering data could also be classified in this category) and six miscellaneous articles dealing with such topics as forecasting techniques, use of telecommunications, linear equation solver etc.

The keynote address 'Automation and telecommunications — the way to improved productivity', shows the need to integrate the technologies involved in automating the factory, design and engineering activities and the office, into a coherent whole, as opposed to the rather piecemeal operations which have characterized the applications of computers etc. in the past. The efficient and controlled flow of data between the various divisions in an organization is an essential part of any such integration. The articles presented in this publication explore and expand current practice and future trends in these areas.

Engineering data

The articles on engineering data handling give a good insight into the large masses of data, both graphical and alphanumeric, which are involved in some large organizations (US Government Departments, aircraft and aerospace industries and other defence contractors etc.), and how technology can be applied to structure and facilitate the use of such information. Many articles concentrate on the problems associated with the generation, release, updating and control of engineering drafting and documentation systems, and some address themselves to the future trends of holding all such data in electronic form, including the work involved in converting the existing mass of such data, in the form of drawings etc., to electronic form. A number of articles deal with the impact of automation technologies on personnel in the drafting and design fields. A most amusing but thought provoking article, entitled 'How to succeed in failure or how not to

implement a CAD system', highlights many of the pitfalls to be avoided in setting up a CAD system.

The articles dealing specifically with computer graphic techniques, rather than the use of graphics or CAD/CAM in existing or projected systems, vary widely, from what appears rather like a company advertising release to some interesting articles on three-dimensional modelling and one which describes a system for creating true three-dimensional objects, which a user can view from different angles by head movement.

In summary, this book, although marred by the organizational deficiencies already described, is a useful acquisition for companies involved, or about to be involved in the CAD/CAM field, since it describes current practices, problems and future systems planned by large organizations who have a great incentive to do the job properly. Although clearly not a text book at either undergraduate or postgraduate level, it should also be of interest to educators in organizations who will produce engineers and computer scientists to enter these areas of engineering planning, design or management.

B C Reardon
University College Dublin, Eire

A course in programming in FORTRAN revised to incorporate FORTRAN 77

V J Calderbank

Chapman and Hall, UK (1983) 183 pp, £5.95

A book bearing the above title promises two things — how to write programs, and how to implement those programs in FORTRAN. This book will teach you nothing about the writing of programs. You may learn the rules of the game, but no winning strategies.

The book was first published in 1969 and shows little awareness of developments in programming methodology since then. The reader will be led to believe that programming is largely a trial-and-error art. The rules of the language are often described by examples or without proper definitions. The programming examples are loosely or incompletely specified, and their purported solutions are presented as *faits accomplis,* without justification. Difficult primitives such as loops are presented, but no means of reasoning with them. Validation of programs is hardly mentioned. Design and implementation are not distinguished. Need we wonder at the poor state of programming when we nurture young students on such books?

Some examples. On page 21 the intrinsic function SQRT is defined as 'finds the square root of a real argument'. I infer that SQRT is complex-valued and that the argument may be negative — but this is surely incorrect. On page 36 a program is presented whose specification is 'to find the sum and the product of the first N integers'. But there is no first integer — perhaps the author intended non-negative or positive integers. Neither is the acceptable range of valus for N stated. Can N be zero? The context would seem to exclude negative values for N, yet the program terminates without error indications for all integral values. On page 38 a program admitting a division by zero is presented. On page 48 a program to illustrate the power of arrays is presented — but the program works equally well if the one array in the program is replaced by a scalar.

Calderbank's book may suit readers who want to know the technical rules of FORTRAN 66, and it is especially well organized for readers wishing to upgrade their FORTRAN 66 rules to FORTRAN 77 rules, the changes in the language are set out in bold type. It will not suit students of FORTRAN 77 who are not also interested in FORTRAN 66.

It may be unfair to single out this book for criticism. Such criticism could be applied to a greater or lesser extent to most textbooks on programming. But the criticism must be made. We will never solve the 'software crisis' until we greatly improve the teaching of elementary programming, no matter what software tools we build or management methods we adopt. In fairness, however, Calderbank's book is probably no worse than many, and cheaper than most.

J Morris
Department of Computer Science, University College Dublin, Eire

Facsimile equipment

W J Welch and P A Wilson

National Computing Centre, UK (1983) 110 pp, £7.00

I read this book as someone intending to buy a facsimile system, and although I haven't actually made the purchase yet, I am quite certain that my £7.00 for the cost of the book is well spent. I found it both informative and easy to follow.

This booklet is part of the National Computing Centre's (NCC) 'Office technology in the '80s' series, and was prepared by the NCC's Office Systems Team. Their stated objectives are to keep potential users informed, to disseminate experiences of larger users and to offer guidance on planning and implementation. The booklet meets this objective, and as it states in its title is an evaluation

book reviews

guide. It is written for a UK user contemplating the purchase of facsimile equipment and gives some useful addresses. It lists potential suppliers and the NCC maintains more up to date information.

The book goes into some detail in listing the criteria for evaluation, dividing them into three main areas: functional criteria, ease of use criteria and suppliers criteria, and then further subidviding each area so that at the lowest level, for example, there is a discussion on what to look out for in the 'paper handling area' or in the 'compatibility' between machines area. Technical criteria are dealt with thoroughly without getting bogged down in too much technical detail. No idea of cost is given in the booklet, and I would have found it very useful to know where the elements of cost were and to have had a feel for the comparative costs of one group of facsimile machines against another, whether a paper-feeding mechanism costs more than a group facsimile machine, or whether the cost of the machine is insignificant compared to telephone charges. I realize these things are changing, but so is the technology, and the life of a guide of this type is short enough to allow some comparative costs to be included.

Having gone through the technical criteria, the booklet then outlines the mechanism of a point system for actually doing the evaluation and gives a set of score sheets.

I shall use this booklet in evaluating my system, and it has been very useful introductory reading to facsimile equipment.

M Sherwood-Smith
Department of Computer Science, University College Dublin, Eire

Viewdata and the information society

M James

Prentice-Hall, USA (1982) 293 pp, £29.95

This is an elegantly produced, readable book, drawing on research material provided by Butler Cox and Partners and LINK/IDC. It is also, unfortunately, already out of date, and contains enough editorial slips and curious statements to cause doubts about the whole (for example, 'Teletex' and 'teletext' are confused in a 'Terminology note' on the title page; 'With dedicated channel teletext . . . this is like having interactive access to 15 000 pages in a machine, with a response time of 0–25 microseconds').

The review of the state of developments in 1981 is comprehensive, and it is pleasant to find a US author doing justice to European initiatives. The various forms of public and private system are well explained, and the alternative roles in information retrieval, small-scale computation, message handling and software distribution are clearly distinguished and described.

This is an expensive book, and since Byte magazine published an excellent 'theme' issue on videotex in July 1983 this would seem a better investment, at least as a starting point for those new to the subject.

J H Ashford
Maidenhead, Berkshire, UK

Text filing and retrieval systems — a practical evaluation guide

S J Newton

National Computing Centre, UK (1983) 130 pp, £9.50

This book is part of the NCC series of evaluation guides concerned with office technology products. It is written for those considering the purchase of a text filing and retrieval system, and is intended to help in the process of choosing one of a range of commercial products, from the early stages of establishing requirements to signing agreements with vendors. The book deals mainly with the user interfaces to such systems and has little evaluation of technical content of filing and retrieval methods.

The book essentially consists of two parts. In the first part, the features of a text filing and retrieval system to be considered when buying such a system are outlined. The criteria to be considered when buying are divided into functional, ease of use and supplier sections. Each of these sections are then further subdivided to give a well structured, well presented and fairly complete picture of filing and retrieval systems. A notable omission from this section, however, is the question of integrating a filing/retrieval system with other office products, such as word processors etc. This is not covered at all.

This first part of the book is intended to make the reader who is unfamiliar with text filing and retrieval systems aware of some of the possible features which could be incorporated into such a system. This section could also be just as useful to system designers or experienced users, in that it gives a global outline of what a system with 'all the frills' could look like.

Having put the reader in a frame of mind such that he should know exactly which features he wants from a text filing and retrieval system, the second part of the book describes a methodology for evaluating commercially available systems. The

evaluation method as described in the book uses pages and pages of evaluation sheets, which are printed as appendices, to determine scores for features. The features to be evaluated are presented as a series of questions, which range from, 'can the user amend the synonym list?' (A7) to, 'is it easy to replace printer ribbons?' (A19). The features are weighted by the reader, normalized according to his relative requirements and the feature scores are added to get a score per 1000 for each system evaluated. This procedure I found to be very long, and I would imagine that most readers would take many short-cuts in their evaluation of commercial systems.

In general, this book is written for a specific purpose (to aid evaluation) and this is does quite adequately. It is not however, a text book, although the first part of the book could be useful to others besides those doing an evaluation.

A F Smeaton
Department of Computer Science, University College Dublin, Eire

Electronic inventions and discoveries

G W A Dummer

Pergamon, UK (1983) 233 pp, £22.00 (hard), £9.95 (soft)

This is a fascinating and very readable compilation, with entries ranging from 1642 (Pascal's computer) to 1982 (Amorphous photosensors). The major part of the work comprises brief histories arranged in strictly chronological order. A valuable part of each entry is the provision of the source of the information and, sometimes, additional references (e.g. to patent literature). However, the book is not merely a catalogue, since it also contains concise histories on a number of aspects of electronic technology: audio, radio and telecommunications, radar, TV, computers with robotics and information technology and industrial applications. These are masterpieces of highly condensed information. Since this book primarily deals with hardware, there is little on essentially software techniques such as programming or information retrieval. As a reference work on the history of electronics this book can be thoroughly recommended.

G L Reed
Department of Computer Science, Brunel University, UK

calendar

Date	*Title*	*Organizers*	*Place*	*Other Details*
1984 25–27 January	8th International Seminar on Security in Banking	INSIG, Institut de formation interbancaire, 40, rue de Monceau, 75008 Paris, France	Paris, France	
13–16 February	3rd kuwait International Information Management Exposition and Conference (Kuwait Info 84)	The Information Centre Company, PO Box 26626, Safat, Kuwait	Safat, Kuwait	Information management systems for Gulf States
13–15 March	Scottish Computer Conference	Contact: Jenny Mann, Quadrilect, Courtfield house, Baldwin Gardens, London EC1N 7SB, UK	Glasgow, Scotland	Microcomputers and microcomputing, communications, electronic office, cost and profit aspects
3–5 April	The International Teleconference Symposium	International Teleconference Symposium, c/o Intelsat, 490 L'Enfant Plaza SW, Washington DC 20024, USA	Philadelphia, USA, London, UK Tokyo, Japan, Sydney, Australia	
9–12 April	IEEE infocom 83	The Institute of Electrical and Electronic Engineers Inc., 345 East 47th Street, New York, NY 10017, USA	Waterloo, Canada	Computer and communications integration: theory and practice
14–17 May	International Conference on Communications: Science-Systems Service	Contact: F T Andrews Esq., Bell Telephone Laboratories, Room 2c 635, Crawfords Corner Road, Holmdel, NJ 07733, USA	Amsterdam, the Netherlands	
14–17 May	4th International Conference on Distributed Computing	Distributed Computing, PO Box 639, Silver Spring, MD 20901, USA	San Francisco, USA	
15–18 May	Communications 84 — International Exhibition of Telecommunications, Radio and Information Technology Products and Services	Publicity Manager (Communications 84), Industrial & Trade Fairs Ltd., Radcliffe House, Blenheim Court, Solihull, West Midlands B91 2BG, UK	Birmingham, UK	
21–25 May	4th Jerusalem Conference on Information Technology	JCIT, PO Box 639, Silver Spring, MD 20901, USA	Jerusalem, Israel	Software engineering and manufacture, economics and management
22–25 May	Infotechasia 84 — 1st South Asian Business, Office Communications, Hardware, Software Show	Overseas Exhibition Services Ltd., 11 Manchester Square, London W1M 5AB, UK	Singapore	
23–25 May	Teleconference and Interactive Media 84	Center for Interactive Programs, Old Radio Hall, 975 Observatory Drive, Madison WI 53706, USA	Madison, USA	
29 July– 2 August	World Conference on Computers in Education — WCCE/85	Chairman, WCCE/85 Organizing Committee, Dept. of Computer Science, Christopher Newport College, Newport News, Virginia 23606, USA	Norfolk, USA	
4–7 September	First International Conference on Human-Computer Interaction (Interact '84)	The Institution of Electrical Engineers, Savoy Place, London WC2R 0BL, UK	London, UK	Human factors in system development; design and evaluation methods; human-computer interface design; impact of computers on human behaviour; human aspects of new horizons
25–30 September	IFIP '84, Network in Office Automation International Symposium	Institute of Mathematics with Computer Center, Symposium Secretariat, IFIP/Network '84, Sofia, 1090, PO Box 373, Bulgaria	Sofia, Bulgaria	Network architecture, microcomputer networks, personal computers in office automation, integrated office systems, reliability, interfaces