Arthur Frederick Parker-Rhodes: a memoir.

Yorick Wilks

Abstract: Parker-Rhodes was an original thinker in information retrieval, quantum mechanics and computational linguistics. The paper describes the basic ideas of his work in the last area for an audience familiar with only the others. It tries to bring out his view of the role of lattices and graphs in language structure and the way his assumptions differed from those of his contemporaries. The paper also aims to give some anecdotal feeling for him and his writings and to point to more recent work where his ideas have reappeared in other forms.

Note: a version of this paper was given as a Parker-Rhodes Memorial Lecture at the annual meeting of the Alternative Natural Philosophy Association in Cambridge, September 1995.

Arthur Frederick Parker Rhodes (AFPR from now on: his colleagues and he all referred to each other by initials at the period I am writing about) was a paradoxical man: never in very good health, he nevertheless dressed as a countryman in Cambridge and walked his collie Shep miles to work every day and back. One says work, but life at the Cambridge Language Research Unit (CLRU) was not like other occupations: as a research student one did not know everything that went on, but we were all pretty sure that, when grant money was short, AFPR was not paid because he was considered to have private means. Later, one learned that, as Communist Party members earlier in their lives, AFPR and Damaris, his wife, had been under constant pressure to hand over their money to the Party, but somehow they never had.

I knew him and the CLRU best during the period 1962-9. It was not part of the University at all and housed an extraordinary collection of eccentrics. It survived, in a way that would now seem almost miraculous, on the instincts and energy of Margaret Masterman its Director (MMB), and the grants she extracted from the UK and US governments and later the EEC Commission. Among those then working there are now numbered an FRS, an FBA and a Stanford full Professor, but also some, no less talented and original, who had no academic careers in the conventional sense, and among whom were AFPR and MMB. AFPR would be modestly pleased that he is now enjoying something of a revival among physicists and that a memorial lecture for him is being given, but not wholly surprised. His surprise might be that it is being given by me!

The work of the CLRU at that time, its golden age probably, was in three parts: research on language and computing, including translation, which was the main interest of MMB; research on information retrieval, by Roger Needham (RMN) and Karen Sparck Jones (KSJ); and the Information Structures Unit, under Ted Bastin (EWB).

Some people had interests in two areas: KSJ in the first two, and Stuart Linney (DSL) in the first and last. AFPR was unique at CLRU in that he actively contributed to all three. He also collaborated with RMN to set out the Theory of Clumps, which has become a standard tool for statistical information retrieval and provides functions with which data that may share no common features can best be clustered into coherent sets or clumps.

My own interests were and are in the first area, language computation, where AFPR is perhaps least known. He did write a book on language (Parker-Rhodes 1978) and I recall that I had something to do with pressing him to get it out and the publishers to take it. But it had all the faults of any AFPR work, the very ones that, along with a lack of ambition, made a conventional academic career impossible for him.

He did, of course meet one essential condition for British academic advancement which is to get to yourself into Oxford, Cambridge or Edinburgh and never, ever, leave.

But that is only, as they used to say, a necessary and not a sufficient condition: he also suffered from an inability to revise a draft, an unwillingness to publish anything he wrote, and a blithe ignorance of the role of citations. The combination of these faults is disastrous. With most of us drafts converge, under criticism or self-criticism, and the last draft is better than the first. AFPR, after criticism, would simply go and write another draft, quite different from

the first, so the process never converged. Also, having written a document--and more about those documents later--his normal procedure was to put in it a drawer, for he had made his point by writing it. He was rather like the mathematician in the old joke who, having proved the raging fire cannot reach his bed, goes back to sleep.

His book had some of these failings which may explain why it fell so stillborn from the press; I had hoped the publishers would sub-edit it and give it the shape he would or could not, but no, they simply printed what he sent without examination, and so were the worst possible combination of author and publisher. As to references and citations, he shared Wittgenstein's view that it was unimportant whether anyone had said the same thing before or not: all that mattered was the content of the thought, which is not a view that endears its holder to colleagues. It was said that he only put some random citations into the book because MMB told him you had to so as to get a book published. Once he saw that was part of the obligatory syntax of book writing, he could carry it out.

Before moving to some detailed discussion of his views on language structure, and how they fitted, or did not, with his other views, one should remember that he did not begin from any assumptions one could call philosophical, in the way that some of his colleagues at CLRU most certainly did see their work as an extension of metaphysics by other means. One could say that AFPR was a Quaker, which may in part explain his distaste for abstract thought until, that is, one remembers that Eddington was one too.

If you asked him if he was hungry he would look at his watch, which was for him a joke about the externalization of sensations. Conversely, he often said, when explaining why he never consulted a dictionary: 'If want to know what a word means, I ask myself'. And he told me once that if he wanted to understand a new subject he wrote a book about it; the remark being in the context of his once having written a history of the world, which I assumed had gone into the bottom drawer. Clearly, anyone who holds the assumptions behind those jokes cannot have a consistent philosophical position: sensations for him were externalized, while at the same time he knew meanings by introspection!

AFPR's original thesis work had been on the statistical growth and development of fungi populations, which had little direct connection with his CLRU interests, except with the statistics in clump theory. But the mycologist in him remained very active: he tramped Nepal with Damaris seeking rare fungi; he loved knowing better than the rest of us what was poisonous and what was not. Tom Sharpe rented a writing hut in the CLRU garden opposite AFPR's, who once said to him, as they looked at a garden fungus, if you believe that to be edible it is surprising there are so few deaths from fungus eating as there are. Although their attitudes to drafting and revision were quite different, the two seem to have got on well.

There was an obsession at CLRU with working in a hut in a garden. I believed at the time it had descended by apostolic succession from Wittgenstein, who had taught MMB briefly, but I could find no reference to hut work in the biographies of him I read later. AFPR had originally worked in the CLRU main building, upstairs and right outside the hidden door to the chapel,

but his thumping on the floor as he thought and typed, humming and moaning, eventually brought about his exile to a hut.

AFPR's intellectual formation was a cluster of diverse interests and influences: Boolean logic was an obvious one, and later lattice theory, to which he was introduced by MMBs interest in it. Michael Halliday, then a lecturer in Chinese at the University and later the major British linguist, introduced AFPR to that language and the notions of formal linguistic syntax, particularly its central notion, present in virtually all syntactic theories, of the way linear items cohere to form longer ones, based on an underlying principle of the substitutability of shorter items for each other, so that some class of larger, including, items remains unchanged with respect to some property, such as "well-formedness".

Here one should add Brouwer's time-centred mathematical intuitionism (Brouwer 1913), which AFPR got from EWB and Clive Kilmister, and perhaps in part from DSL, who knew intuitionistic logic, as well as Spencer-Brown's idiosyncratic notion of logic (1969) based on no more than the ability to make a distinction of this from that (with Brouwer it had been present from past). These notions were important to him, I think, in both his physics and linguistic work.

Associated with CLRU at that period were also cyberneticians, a word now rarely used, but one accepted by Gordon Pask and Robin McKinnon Wood, who collaborated with CLRU at the time. The cybernetic notion of self-organisation was crucial to the ISU approach to physics, one quite different from the then dominant artificial intelligence paradigm. But AFPR was never a cybernetician at all, for that always involved a process view of the world, using continuous mathematics. AFPR wanted a structural theory, using discrete mathematics, at least where language and physics were concerned.

AFPR's writings and thoughts on language processing and structure were not all of a piece. Some were occasional items, such as his astonishing 'A constructive criticism of NUDE' (1965), a title which requires deconstruction, as some would say. NUDE was a semi-formal language (from 'naked ideas') based on ideas first set out by R.H. Richens, a biologist interested in plant classification. It had about eighty primitive items, Anglo-Saxon monosyllables for the most part (like FOLK, STUFF, MAN, THING), along with brackets and connectives. The idea was to use NUDE formulae to describe the meaning of sentences in a way that was formal, removed from English or any other natural language, but which could be used by a program to, say, translate from one language to another by means of an intermediate language, or interlingua, such as NUDE.

Although eccentric, NUDE came to look strikingly like the languages for semantic coding derived much later by linguists for meaning representation (e.g. Jackendoff, 1990). But AFPR was never very warm towards the interlingual stream in CLRU thought, which was basically MMBs. When accused by her of scepticism, his practice was to sit down at his typewriter and immediately write what he considered a constructive contribution. That typewriter deserves a special mention: AFPR did not believe in advanced machines, even those of the Sixties. He had an old fashioned manual typewriter, with which he was able to create mathematical symbols at will by advancing the carriage half a space and then overtyping another letter, a process that

produced a wholly novel and extensible alphabet that looked a little like Hebrew. In the same spirit, he had a beautiful hand calculator that he used with astonishing versatility long after there were tolerable electronic ones.

Yet the AFPR paper I refer to is extraordinary: he noticed immediately, what still is not obvious to linguists and logicians, that it is decidedly odd to code a natural language in forms that are plainly those of a real language (English in capital letters in the case of NUDE). So he set out to construct an artificial interlingua, with a vocabulary of about 150 primitive items, and a syntax more rigorous than that of NUDE but also speakable, so that its 'sentences' were no longer than what they encoded: for NUDE formulae could easily be twenty times the length of the word they expressed.

Here is a part of his language, which I think never escaped from AFPR's drawer, except for a draft he gave me. It was, perhaps, too clever and too ingenious to have been adopted, even if published. It might conceivably have become a new Esperanto for people who wanted a smaller language, easier to learn; AFPR remained fascinated by artificial languages and that must have been part of his motivation in constructing this language. He also imposed a very odd and original constraint on his language, which he believed would make it transparent, in the sense that two observers would always agree on what an expression meant (which was not always the case with NUDE).

He postulated that every string of syllables could be a constituent of the language. On the face of it this is an odd demand, and one not met by any actual language. It would mean, for example, that if English were a language obeying such a constraint, then in 'I want to go to Los Angeles and be in movies' the strings such as (Angeles and) or (go to Los) would all be constituents of the sentence, which they are plainly not, in the sense that (conventionally) 'I' is its subject constituent and 'want to go' its complex main verb and 'Los Angeles' a proper name within it. One could say that the whole purpose of the syntax of a language is to give a formal device to specify which substrings of a sentence are constituents and which are not.

On further examination, however, this conventional wisdom turns out to rest on highly intuitive assumptions about linguistic propriety: Steedman (1982) proposed later that it could make perfect sense to postulate a parser for English in which every string, of any length from the beginning of a sentence, could be considered a constituent: a claim that violates exactly the deep intuitions that AFPR's principle does. However, AFPR was not making a claim about English, but stating a formation rule for a language he was inventing. In practice, and from the fragment of the language he gives, his rule would have led to great difficulties. One could say that it violates the whole notion of a "higher order" item if any lower order selection can be one: that it makes such a language informationally inexpressive, in the way that a language would be if any string from its alphabet was a word. We would then lose redundancy if a character was lost in transmission, since what was left would always still be a word.

AFPRs suggestion may or may not be open to a similar objection, an issue which is closed since the paper was effectively buried and had no successors or discussion, and I mention it only to bring out his ingenuity and playful creativity: a whole new speakable mini-language was created overnight to answer a charge of unhelpfulness about NUDE. A point worth drawing out here in passing is that AFPR, unlike most contemporary linguists, really did know languages, at least as regards reading them. He learned them as a hobby and would sometimes admit to twenty three: he said they became easier after the first half-dozen because you probably had a hook to hang a new one on.

Inferential Semantics

AFPR's major statement on language structure is the book Inferential Semantics, which has not been influential, in spite of its visionary quality. It begins with his own special kind of gusto: a brief excursion into the notion of self-reference which is implicitly illustrated by the fact that the book itself is listed in its own bibliography. A footnote warns us that this is not normal practice.

It was not his first theory of language structure, which was probably the method for syntactic parsing he developed from Michael Halliday's grammatical theories (1965) and which I attempted to program on a Hollerith punch card machine in about 1962-3. His method was to map Halliday's syntactic constituents for a sentence, and their inclusion relations, onto some elementary Boolean lattice and then to attempt to determine the correct parsing by lattice operations. In fact, the punch card machine was a clumsy but perfect

engine for the purpose because the computation involved none of the elaborate binary codings common at that time: the physical overlap of punched holes from card to card determined the set intersection relations and the results could have been read by eye.

He always attributed his use of lattices for natural language structure to MMB, although her original insight was quite different from what he came to use those structures for. MMB believed that a thesaurus, like Roget's, which classified words by their meanings, by placing them under about 1000 general head concepts (MOTION, SUBSTANCE, VEHICLES etc.) was a key tool for the computation of meaning, and of more use than dictionaries, because a thesaurus worked from the meaning (the head concept) to the word and not the other way about (as a dictionary could be said to do).

I will assume the general notion of a lattice as a system of partially ordered sets such that every pair of such sets has a unique upper and lower bound, usually taken to be Boolean set intersection and union. If the generators of the lattice (in the graph representation, these are conventionally the line of nodes drawn horizontally across the middle) are identified with sets corresponding to the 1000 head concepts of Roget then, if a thesaurus is a lattice, one will expect each pair of heads (and so the sets of words associated with them) to have a unique intersection and union, and each of those again to have a unique intersection and union with any other set in the system. In that way overlaps of word sets under thesaurus heads can be calculated in a definite way, and this, MMB believed, could form the basis of calculations to, say, resolve semantic ambiguity in texts as part of a machine translation project.

That particular method was not wholly successful when experiments were done, and MMB did not always distinguish between assuming a thesaurus was represented by a lattice and showing that it was. Nonetheless, the idea was well ahead of its time in terms of using an existing language data base for natural language processing, an idea that only became practical in the last ten years. Even interest in the use of Roget has revived and Yarowsky (1991), using a complex refinement of MMBs original method, now claims to resolve over 98% of words in text successfully to a unique sense, where that is taken to mean attaching a word in context to one and only one thesaurus head.

The main motivation for using a lattice was one of multiple classification: the desire to put a word context in a number of sets simultaneously, in a way one cannot do with trees, then the dominant linguistic data structure. AFPR, though, was never very committed to MMBs main idea of the lexical lattice, a lattice of words classified by their senses: in his book he gives clear counter examples to this as a general structure for lexicons, yet, nonetheless, he continues throughout to use it as an ideal lexical structure. This is almost certainly the first work to give a full abstract structure for an ideal lexicon (apart from MMB's own work), an idea that has revived in recent times with inheritance structures like Gazdar and Evans' DATR (1990).

AFPRs goal in the book, as its title suggests, was not primarily lexical but a general theory of inference in natural language and, most ambitiously, the relationship of inference to understanding. The lattice gave, as does no other plausible elementary set theoretic structure, the closure of inference that he wanted. It is an old chestnut in the philosophy of mathematics that proof and understanding are related: that, in mathematics more than ordinary language,

one understands a proposition by proving it, and the proof becomes a context that gives the theorem its interpretation. This notion had been extended to natural language by Wittgenstein and in particular by R. Bosanquet in the 1940s whose work was published after his death (1945) by Richard Braithwaite, MMBs husband, so AFPR could have known of it.

My own thesis work, to which AFPR's book makes reference (1968), constructed a computational model of that idea, that understanding and a very weak notion of inference in text were closely related. But AFPR wanted a more radical relationship: his basic claim in the book is that there is a weak notion of inference a --> b just when understanding a allows b to be understood. His inference lattices are a model of that relationship, and provided wider textual context by means of binary, applications of lattice operations. I find the fundamental notion fascinating but unbelievable, in that, given any sentence I understand I could not then tell you much about what set of sentences it allows me to understand. In the simple manner of standard logic, one can say that:

I have a red hat

allows me to deduce just

I have a hat

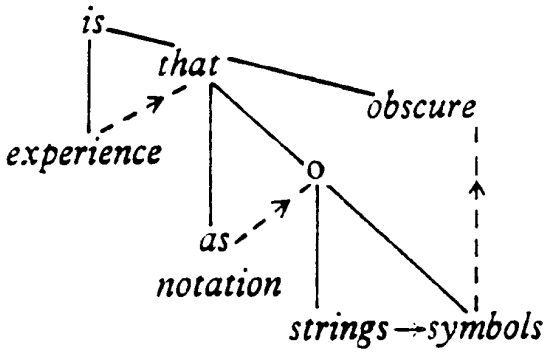
and

I have something

but I cannot be sure that understanding the first allows me to understand the last because I may not happen to know the word 'something'; perhaps I know no words over four letters long.

With this basic relationship defined, however, AFPR's game is afoot: the lattice gives the closure of the structure of inferences of a language (See Figure 2 below) and the sets at the nodes of the lattice are structures representing rhemas, the utterances themselves, whose structures are directed graphs (See Figure 3 below). He says he decided, wisely in my view, against calling the book simply Rhematics. The basic mathematics are then all from either Birkhoff's Lattice Theory (1949) or Harary's Graph Theory (1969) with some original bits to tie the two together via node algebras, so that an associated set of sentences at a lattice node correspond to a set of graphs that themselves form a sublattice of the whole.

The following example will illustrate the three notational methods described: it is based on the utterance "*It is commonly found that symbol-strings are notationally obscure*", which I interpret rhematically as:



. . . 5.38.1

is	**d**	**d**	a	a	a	a	a
experience	a	a	a	a	a	a	
that	**d**	**d**	a	a	a	**d**	
as notation	a	a	a	a			
Ø	**d**	**d**	a				
symbols	a	a					
strings	a						
obscure							

. . . 5.38.2

(is): 01 / (experience): / (that): 014 / (as notation): / (): 01/ (symbols): / (strings): / (obscure):. . . . 5.38.3

In the diagrammatic notation, the designatory arcs are indicated in solid lines, the accessory arcs, where they do not coincide these, as dotted lines (with arrows, since they sometimes have to run upwards, whereas the designatory arcs can always be drawn running downwards).

Figure 2: AFPR's rhemas (p.105).

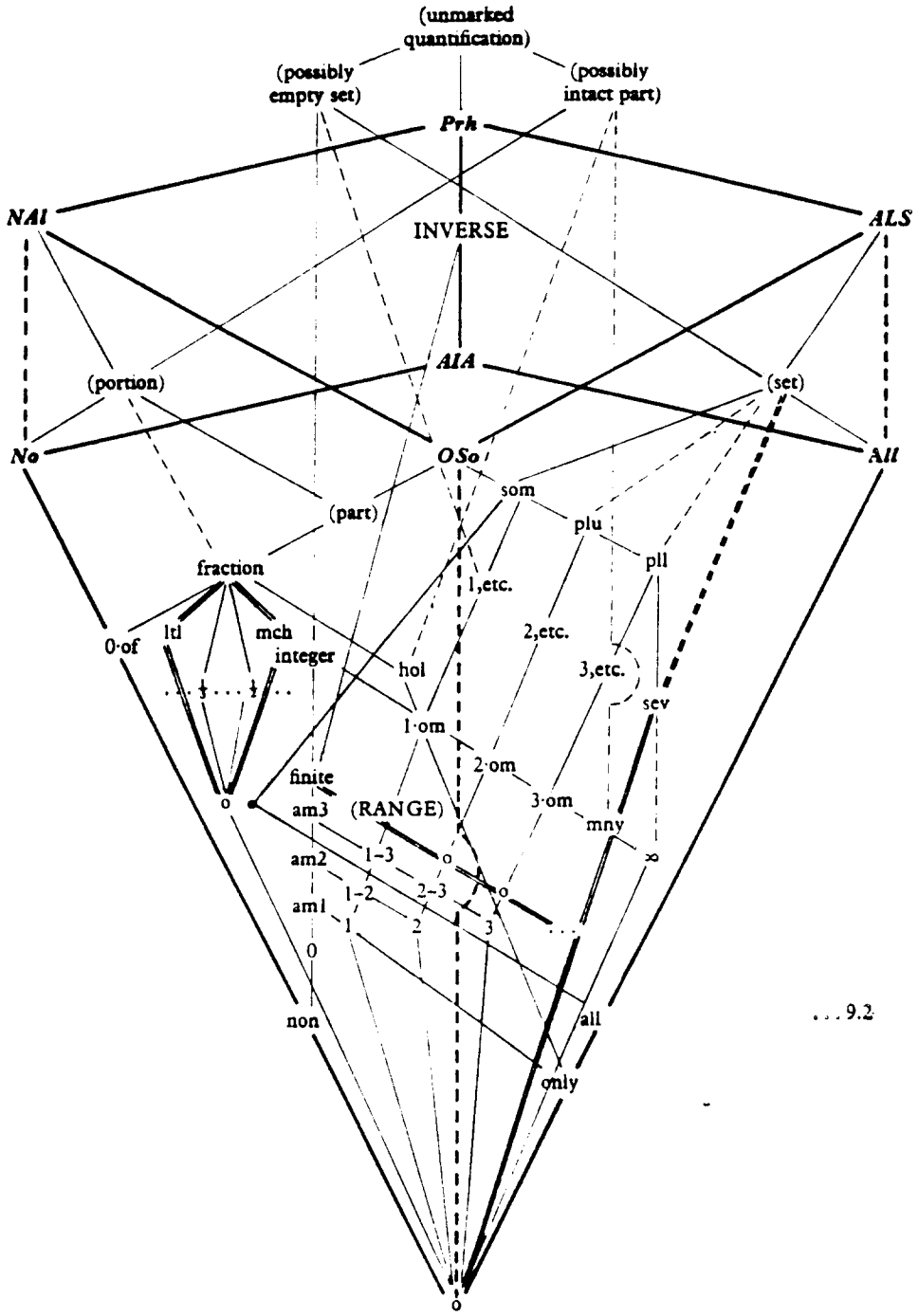


Figure 3: The Quantification Lattice

Into the rhema graphs, AFPR packs all the then fashionable parts of fringe linguistics: such as case grammar and speech acts, for graph theory is infinitely flexible. The originality is the uniform structural treatment of all levels of phenomena by a method utterly distinct from then conventional linguistics which had (and has) few structures beyond trees and a little elementary logic. But AFPR demands more than such a structure can offer; he claims at one point (ibid. p. 63) that the structure yields decidability, which means an algorithm to decide whether or not one can understand some b given some a. I cannot believe that such a fragile relationship can be captured by so powerful a machine.

Notice there is no role for process here - it is all abstract structures - nor is there any role for data, which is now the fashion in linguistics again, and AFPR strongly rejects any form of empiricism. I believe that, strong as his vision was, data could not be provided for it in any straightforward way since, as I asked above, how could one find, in a determinate manner, the set of sentences that an understood sentence allowed one to understand? And, just as AFPR's notion of structure is far more powerful than anything in linguistics then or since, it is not in any sense artificial intelligence either, since it accords no role to the knowledge required for language understanding: for him, understanding is simply a primitive relationship on which everything else rests.

AFPR had no time for Chomsky, normally considered the founder of the formal turn in post-war linguistics, yet they have much in common: the main difference between them being that Chomsky was satisfied with a much simpler mathematical structure for language, that of general logic itself, in Post's production rule form. Chomsky, too, worried about the decidability of his system, transformational grammar: it was not decidable, as became clear. Like AFPR, he had no real interest in computation or processes, though he did, initially at least, believe in structures of simple propositional meanings, which he called the deep structures underlying language utterances, a notion AFPR could never accept.

In the end, I think the most original idea in AFPR's work is the implausible primitive notion of understanding on which everything rests: it becomes somewhat more acceptable when one realises it has nothing at all to do with inference as normally understood, but only with the relationship between what he calls 'repertory' (the set of rhemas inferable from a rhema, in his special sense of infer) and 'rehearsal' (the set of surface rhemas that make up a particular discourse). It is this relationship that the whole magnificent formal architecture is designed to capture and behind it all is his conviction that understanding/inference is a more local matter than normally thought, where by "normal" I mean those aspects of contemporary continental thought, as well as Wittgenstein a generation earlier, that have conditioned us to believe that to understand a sentence fully we have to understand a whole language, and that everything is part of everything else.

But there is a process-related and cognitive strand of thought abroad these days, well supported by evidence, (e.g. Walker and Rambow, 1994) that shows the locality and finiteness of normal conversational inferences, and simple graphical structures have recently been suggested for capturing them. AFPR's architectonic could be seen as an early model for that, and for local

closure, one that kept the inferential repertory very close to the surface rehearsal of language.

Representation and Computation

An obvious feature of AFPR's language work, published and unpublished, was that he freely used strings of binary numbers not just as representational devices for language constituents, in the sense of naming them by binary numbers rather than by conventional forms like NOUN-PHRASE or ANIMATE-NOUN, but also as contentful strings on which direct computation could be done.

This view marked him off totally from contemporary language work in both linguistics and artificial intelligence, where it had been an item of dogma since the mid-Sixties that:

(1) arbitrary names were absurd and a way of deceiving ourselves about what we knew about language, and that

(2) although computers, and perhaps brains, worked with binary strings, that level of representation had no significance, in that it was merely a machine code version of the higher structures with which brains and computers really functioned.

That 'higher' level was what AI languages like LISP and prolog manipulated best, or which could even, according to some, be a form of English-like brain language, or indeed the plainly English-like forms of NUDE and contemporary linguistic representations. AFPR could not accept this: as we saw with his critique of NUDE, he actually wanted arbitrary representations, which he thought users should take the trouble to learn, implicitly rejecting the riposte that those arbitrary languages would then, on being learned, have all the ambiguity of actual languages, as Esperanto itself is said to have developed.

Although one could reject his level of representations as proper representations, nevertheless AFPR did remain committed to representations as such, and to simple mathematical structures like lattices and Boolean logic for manipulating them. Had he survived to see today's swing of the pendulum in AI back towards statistical methods in the form of connectionism and neural nets, AFPR might have been expected to welcome a movement that was anti-representational and to have found structures like binary vectors a proper starting structure for a language 'acquisition' device, and to have reused statistical functions which, while not those of his clump theory, were not a million miles away either.

However, I suspect he would not in fact have welcomed that development, partly because he believed low-level representations accessible as such, and interpretable, perhaps even at the brain level, and because he did not seek an overall consistency between the phenomena he treated statistically (such as clumps over data sets) and those he thought of combinatorially (such as language and quantum phenomena). However, his addiction to binary code---which I have argued was more than the conventional attachment of a programmer to much earlier generations of machines------and his desire to

avoid the self-representation of a language by moving to a wholly arbitrary interlingual scheme, showed that he could not accept the main assumption of those in the so-called Language of Thought movement do accept. That is that language can, in brain or computer, be its own representation and models of it are not like models in science where the model and modelled are, by definition, different types of formal object. His book began, you will recall, with warnings against the perils of self-reference.

I close with this point because, so far as I understand it, that self-modelling is exactly what he did come to accept in the realm of physics where, like von Weizaecker (1971), he was 'against models as such' for quantum phenomena since his combinatorial logic itself became the object of study (as opposed to being modelled by it, as in the conventional view of science). If that is the case, AFPR again did not seek any kind of metaphysical consistency between the study of language and quanta. And why should he, as they are plainly not the same? Of course, since I happen to hold the "self-description" view of language as possible and indeed necessary, I wish AFPR had held it, too, but there it is. Things might well have turned out differently if he had had someone to talk to about his language ideas at CLRU as he did in information retrieval and quantum physics.

Let me leave you with his words, not mine: he once returned to work after a friend's funeral and on being asked, conventionally, how it went, he said, 'Oh, it was quite successful; yes, very successful'.

References

Birkhoff, G. (1949) Lattice Theory. Monographs of the Amer. Assn. of Mathematics.

Bosanquet, R. (1945) Remarks on Spinoza's Ethics. Mind.

Brouwer, L.E.J. (1913) Intuitionism. Bull. Am. Math. Soc. Vol XX.

Chomsky, N. (1966) Aspects of the Theory of Syntax, MIT Press.

Evans, R. and Gazdar, G. (1990) The DATR Papers, Volume I. University of Sussex, Cognitive Science Research Reports.

Halliday, M. (1965) Categories of the theory of grammar. Word.

Harary, F. (1969) Graph Theory. Addison-Wesley.

Jackendoff, R. (1990) Semantic Structures. MIT Press: Cambridge, MA.

Masterman, M. (1956) Potentialities of a Mechanical Thesaurus. Proc. Internat. Conf. Mech. Trans., MIT Press. (in Machine Translation. Vol. 3).

Parker-Rhodes, A. F. (1965?) A constructive criticism of "NUDE". unpublished ms.

Parker-Rhodes, A. F. (1978) Inferential Semantics. Harvester: Hassocks.

Spencer-Brown, G. (1969) The Laws of Form. Allenx Unwin: London.

Steedman, M. (1982) Lecture at the University of Essex.

von Weizaecker, F. (1971) The unity of physics, In Bastin (ed.)
Quantum Theory and Beyond. Cambridge: Cambridge University Press.

Walker, M. and O. Rambow. (1994) The role of cognitive modelling in achieving communicative intentions. In Proc. 7th. Internat. Workshop on Natural Language Generation, Kennebunkport, ME.

Wilks, Y. (1968) Argument and Proof. Cambridge University Ph. D. thesis.

Wilks, Y. (1971) Decidability and Natural Language. Mind.

Yarowsky, D. (1991) Word-sense disambiguation using statistical models of Roget's categories, trained on very large corpora. Proc. COLING92, Nantes.