Hospitable Formalism

R. A. Fairthorne

The formalism of a developing science has to grow with, indeed rather ahead of, that science, adapting to changes in the phenomena studied and the way in which they are studied and interpreted. The assumption behind all formalized sciences is that this is possible, if only in retrospect, provided that changes in interpretation do not contradict experience as it stands at the time. It is also assumed by many that where the existing formalism is inadequate to novel situations, the inadequacy arises from the formalism being a special case of a wider or deeper unifying formalism. This last assumption implies that even if drastic changes take place in the understanding of the phenomena, these affect techniques and procedures only in the novel areas. For instance, even after many years relativistic principles are still irrelevant to most applications of physics to engineering, and where relevant do not invalidate the techniques of Newton, d'Alembert and St. Venant, but extend them. On the other hand, understanding and interpretation of mechanical phenomena is profoundly affected.

This suggests that for a developing science the soundest approach is to start formalism at the most concrete procedural level that just qualifies as a part of information science, and then develop outwards and upwards. If the formalism is on the correct lines its elements, structures and principles will need redefinition and refinement as its scope and store of experience expands but, with good fortune and some prophetic talent, these modifications will be ahead of existing techniques for dealing with current situations. They will, in fact, allow early warning of changes in technique through scale factors that indicate the balance of influences on phenomena (cf the Reynolds, Mach and Glauert Numbers in aeronautics). To

achieve such scaling factors we must to some extent know the answers in advance; one of the less advertised requirements for all science, technology and mathematics. But even apart from this there are two obvious springboards, Information (ie Signalling) Theory, and Discourse in the timeless, non-spatial sense of, say, library classification. The first contains the distinct, if not strictly defined, elements of Message, Code and Channel. A message here is an abstract entity, one of a socially agreed set of distinct items. The task of the signaller is to indicate which message has been chosen by means of a set of patterns, the code, that have to be represented by patterns of events in a physical medium and environment, the channel.

Discourse, in the narrowest effective sense, is not directly concerned with the physical world, it deals directly with whom, the Source, speaks to whom, the Destination, about what, the Designation.

Neither Signalling nor Discourse can exist without the other, or without other activities. But between them they certainly contain the six elements that are enough to formalize basic procedures, provided that we look at the latter as of the form 'influence of A upon B as mediated by C'. The twenty such configurations cover all the procedures of *Notification*, ie those parts of discourse that are delegable, and are not changed by the discourse. That is, Notification is the frozen, 'snapshot', aspect of discourse at a given time and place (the 'timelessness' and 'unlocalized' nature of discourse exists only at the place of the delegable activity and only as long as the activity lasts).

If we study the successive notification structures as units of a higher level we should arrive at a formalism of discourse in which the structure of discourse is changed by reason of the discourse itself, the level of Conversation. The procedures of Conversation, whatever they may turn out to be (clues to this lie in Rhetoric), should in their turn be the elements of the next higher level, in which not only the terms of discourse, but also the cognitive structure of the

discussants changes by reason of the discourse. This is the level, or floor of the level, of Cognition.

Such an approach, by elements and levels, is hospitable to various philosophical viewpoints. For instance, on the Notification level, both the view that science 'breaks the code' of the physical universe, and that it constructs a code to fit observation, can be accommodated though not, of course, at the same time. The first view is represented by conjoining the Observation and Readership triads, the second by conjoining Observation and Authorship.

A possibly greater advantage is that it not only clears up confusions between various interpretations of 'information' and its measures, but also indicates where different interpretations and measures (where applicable) are likely to be found. The selective information of Shannon clearly applies on the Notification level, in the sense that similar entropic measures - not necessarily of any kind of information in any sense - can be applied to any of the twenty triads, not only to Signalling. Information about changes of the structure of discourse can be measured only at the second level, Conversation. Its measure may have the same entropic form, but the elements will no longer be the elements of Notification but the procedures of Notification.

Similarly, for the various manifestations of 'Information' at the Cognitive level.

The mathematical tools of information science should be equally hospitable in order to keep pace with the development of the formalism. Statistical and combinatorial techniques, properly selected, do not need any boosting. Algebraic techniques as usually employed are inadequate to deal with even the existing situations. This could be corrected if Boolean algebra applied to Notification procedures were regarded as a degenerate form of Brouwerian or its dual, Heyting algebra, eg even if the difference between 'all-but-not-only' and 'only-but-not-all' parts of a collection were vanishingly small, there would still be a limit on the length of a hierarchical

chain and on the number of terms that can be effectively co-ordinated in retrieval operations. Similar considerations give limits to the precision of (selective) information measures corresponding to these situations. The parallel here is to the effect of even a vanishingly small viscosity in the formulation of fluid flow. Introduction of a zone of imprecision, however narrow, between a subset and its complement, (which is, in terms of information retrieval, the distinction between Brouwerian and Boolean algebras) formally extends the techniques to cover the joint effects, but may have no numerically significant effect throughout a given operation at a given place and time.

---oOo---