Meta- and Object-language in Information Retrieval Research

Julian Warner

The Queen's University of Belfast

Visiting Scholar, The University of Illinois at Urbana-Champaign

9 February 2000 – 14 September 2000

Abstract

The distinction between a meta- and an object-language has become increasingly familiar in information science, through the diffusion of the concept of metadata. A significant antecedent can be found in the development of formal logic. This paper proposes an analogous distinction for information retrieval research, between discourse about and transformations within systems. The aim is to obtain analytical clarity and economy in research effort.

1. Introduction

The distinction between a meta- and an object-language is familiar from the contrast between metadata and the objects described (although the term object-data is not necessarily used). Antecedents in information retrieval research can be found in analyses of indexing vocabulary, where the metalanguage of index terms (either derived from a thesaurus or classification scheme or generated from the object-language) is differentiated from the object-language described. The distinction is valuable for revealing commonalities between apparently different approaches to indexing (Gardin 1968). Fuller antecedents can be discovered in formal logic where a distinction between the verbal metalanguage of discussion and interpretation and the symbolic object-language of a logical system has been developed and found to yield analytical clarity (Bochenski 1961). This paper reviews the development of the distinction in formal logic, providing a historical context for current discussions of metadata, and then

applies an analogous distinction to information retrieval research. A significant gain in analytical clarity, which can inform future developments, is obtained.

2. Meta- and object-language

The distinction between meta- and object-language in formal logic is unlikely to develop while undifferentiated verbal forms are used both for interpretation and discussion and for the logical formulas interpreted. Following the increased development and use of logical notations from the mid-19th century, distinctions were formulated in the late 19th and early 20th centuries. The distinction has been generally, although not universally, accepted in formal logic and is recognized to have a simplifying effect (Bochenski 1961), although it has been subject to different interpretations. In this context, it is understood as differentiating a verbal metalanguage from the symbolic object-language it speaks about, following Ramsey (1990). The concern here is not directly with the final validity of the distinction but with the example of its clarifying effect.

The metalanguage of discussion about information retrieval systems can be differentiated from the object-language of the transformations carried out within information retrieval systems. The distinction between meta- and object-language can be discovered in information retrieval research but has not been fully known to it.

The metalanguage of information retrieval research has characteristically, although not necessarily explicitly, displayed a founding assumption, some central concepts subsidiary to that assumption, entities distinguished for evaluative purposes, and some derived measures. The founding assumption has been of the value of delivering all, and possibly only all, the records (or other objects) relevant to an information need; the central concepts information need and relevance; entities differentiated have included the query, or, in an alternative formulation, the assertion of relevance; and derived measures precision and recall, although others have been proposed. Information retrieval research discussion has also been concerned with the relative value of different transformations in the object language when assessed within this paradigm. The

potential transformations have been revealed on analysis to be variations on primitive operations of sorting or partitioning and the transformation of one symbol into another (Buckland and Plaunt 1994). From the perspective here, this can be a regarded as a special case of the known potential for reducing mathematical and logical operations in the object-language to the writing, erasure and substitution of symbols (Ramsey 1990). The founding assumption of information retrieval research (of the value of delivering relevant records) has seldom been explored, although there has been greater questioning of its central concepts (information need and relevance) and derived measures (precision and recall) (Ellis 1984; 1996). Alternatives in the object-language (for instance, non-Boolean searching) have also been elevated to the status of ends rather than regarded as techniques.

Pursuing the distinction between meta- and object-language can yield a simpler metalanguage while preserving the intellectual labor embodied in the object-language. Within the metalanguage, a verbally articulated query and reduction of relevance to binary or quantified judgements have already been revealed as intra-theoretic, connected with the demands of experimental evaluation, not intrinsic to searching information systems (Ellis 1984; Heine 1980). A further development has been to replace the founding assumption of the value of delivering relevant records with an emphasis on enhancing the capacity for informed choice (Warner 2000). This broadly corresponds to the exploratory capability also advocated as an evaluative criterion (Ellis 1984) but the antithesis with relevance is dissolved. Relevance is regarded as a significant dimension of informed choice or exploratory capability. Techniques developed in information retrieval research for transformations in the object-language can be adapted, perhaps particularly in interactive modes, to support exploration rather than being used to deliver relevant records. The intellectual labor embodied in one component (the object-language) of information retrieval research has been preserved while its founding assumption has been rejected. Other components of its metalanguage, such as the central concepts of information need and

relevance and the derived measures of precision and recall, can be transformed into objects of dialectic engagement for future development.

3. Conclusion

In conclusion, the application of the distinction between meta- and object-language to information retrieval research has been found to give a significant gain in analytical clarity. What is regarded as valuable in information retrieval research has been retained. A continuing dialogue between theory and practice is also enabled (Warner 2000). The substitution of the principle of enhanced informed choice can be regarded as exemplifying the development of a true science, in which previous paradigms are absorbed into a new formulation as special cases.

References

Bochenski, I.M. 1961. *History of formal logic*. Translated by I. Thomas. Indiana: Notre Dame, 1961.

Buckland, Michael and Plaunt, Christian. 1994. On the construction of selection systems. *Library hi tech* 12: 15-28.

Ellis, David. 1984. Theory and explanation in information retrieval research. *Journal of information science* 8: 25-38.

Ellis, David. 1996. *Progress and problems in information retrieval*. London: Library Association, 1996.

Gardin, Jean-Claude. 1968. Document analysis and linguistic theory. *Journal of Documentation* 29: 137-168.

Heine, Michael H. The 'question' as a fundamental variable in information science. In O. Harbo and L. Kajberg (eds.). *Theory and application of information research* (Proceedings of the second international research forum on information science, 3-6 August 1977, Royal School of Librarianship, Copenhagen). London: Mansell, 1980. 137-145.

Ramsey, Frank Plumpton. 1990. The foundations of mathematics. In: D. H. Mellor (ed.). *Philosophical papers*. Cambridge etc.: Cambridge University Press, 1990. 164-224.

Warner, Julian. 2000. In the catalogue ye go for men: evaluation criteria for information retrieval systems. *Aslib Proceedings* 52: 76-82. And: Information Research, Volume 4 No. 4 June 1999. At:

http://www.shef.ac.uk/~is/publications/infres/paper62.html.