Chapter 8

Conclusions and recommendations

8.1. Conclusion

In this project, we set out to design and construct a front-end
system which would enable weighted searching with relevance feedback to
be carried out on a conventional Boolean host. In this endeavour we
were successful, and the resulting system has been described in these
pages.

The system certainly has limitations, and requires development in
certain directions. Some of the limitations are of a fundamental
nature, and may not be removable within the present framework (for exam-
ple, it seems likely that any algorithm that was used for the searching
would be limited to a relatively small number of terms). Nevertheless,
it would appear that the system, with a little further development,
could form the basis of the sort of evaluation project originally
envisaged for the Oddy-Jamieson system.

8.2. Oddy-Jamieson evaluation

The proposal was to install the intelligent terminal/front-end sys-
tem within an operational environment, and use it as an alternative to a
conventional dumb terminal in a controlled experiment (but with real
users). The operational environment chosen was the Central Information
Service at the University of London (Jamieson and Oddy, 1979).

Oddy and Jamieson worked out an experimental plan in some detail.
It involved randomly assigning users to either the experimental system
or the conventional one, and using various kinds of post-search assess-
ment. In essence it remains a good plan, although some of the practical
details would need to be changed.

8.3. Possible operational evaluation of cirt

The first question that arises concerns the mode of implementation
of cirt in the operational environment. In our original proposal,
several options were outlined. The major distinction between them con-
cerned whether to install cirt physically in the new location, or alter-
natively to provide access to it from a dumb terminal at the new loca-
tion.

The balance of advantage at present seems to us to come down
heavily on the side of keeping cirt where it is at the moment and pro-
viding access through the network. For one thing, it is clear that in
such a project, the job of maintaining cirt would be quite a substantial
one: one would have to keep a close watch on how it runs under opera-
tional conditions, in order to detect and remove any remaining technical
difficulties. (Our experience leads us to assume that there will be
further technical difficulties, either with cirt itself or with the

network software.) A second and related point is that it may no longer be possible to get enough searches for the evaluation at CIS itself, since University of London searches are more and more being done in the Institutions themselves. Thus one might have to conduct the experiment usng searches from several different locations. Installing (and maintaining) cirt in several locations would not seem to be a desirable course of action at the present stage of development.

Such arguments do not of course preclude the eventual development of a portable system; indeed, we would regard that as a long-term aim.

If we were to pursue the course suggested here, the various Institutes of the University of London would very probably be able to get access to cirt through ULCC and Joint Academic Network connections. These connections are at present being installed; there may be some delay before all the necessary connections are available, and clearly any new evaluation project will have to wait until they are.

8.4. Outline of a project

It would appear appropriate, therefore, to consider a project along the following lines:

a. The main aim of the project would be to conduct a controlled experiment, in an operational environment, to compare weighted-searching-with-relevance-feedback with a more conventional approach.

b. There would need to be an initial stage of further development of cirt. The main requirements are: accounting procedures; some development of the user interface; any work required to ensure that any new network connections are operating correctly. This initial stage would require the employment of a programmer, and would last a few months (3-6).

c. The second stage would be the evaluation, as envisaged by Oddy and Jamieson, but perhaps involving other University of London institutions as well as CIS. This would last about one year, and would involve employing an information scientist, to be based at CIS. The programmer would continue on the project, to maintain the system, provide any necessary technical support, and (in her/his spare time) work on some of the developments of cirt suggested in this report.

d. Thus overall the project would require about two and a half man-years, and money for data-base searching, spread over one and a half calendar years. Because of the requirements for network connections, it could not start before October 1984 at the earliest.

References

Croft, W.B. and Harper, D.J. (1979) Using probabilistic models of document retrieval without relevance information. Journal of Documentation, **35**, 285-295.

Harper, D.J. (1980) Relevance feedback in document retrieval. Ph.D. Thesis, University of Cambridge.

Jamieson, S.H. (1979) The economic implementation of experimental retrieval techniques on a very large scale using an intelligent terminal. In: Proceedings of the Second International Conference on Information Storage and Retrieval, Dallas, 1979. New York: Association for Computing Machinery, 45-51.

Jamieson, S.H. and Oddy, R.N. (1979) Implementation and evaluation of interactive retrieval through an intelligent terminal. Project proposal to the British Library Research and Development Department.

Morrissey, J. (1981) An intelligent terminal to access Euronet. Student project, Computer Science Department, University College Dublin.

Robertson, S.E. and Sparck Jones, K. (1976) Relevance weighting of search terms. Journal of the American Society for Information Science, 27, 129-146.

Sparck Jones, K. (1972) A statistical interpretation of term specificity and its application in retrieval. Journal of Documentation, 28, 11-21.

Weiss, S.F. (1981) A probabilistic algorithm for nearest neighbour searching. In: Information Retrieval Research. London: Butterworths.