Types of test

Part 2 is devoted to different kinds of retrieval system test, showing how the general considerations of Part 1 apply in types of test context. Thus direct tests of retrieval systems may, as already indicated, be either investigations, typically of single systems, or experiments with varying degrees of control implying a range of systems; and they may be tests in operational or laboratory environments. Experiments impose more constraints on test designs than investigations, so the fact that more information may be got from them has to be balanced by the fact that system operation under these constraints may not be natural. Again, there are important differences between studies of the operation of real systems with real users, where tests are not completely repeatable, and detached laboratory studies, where assumptions about the behaviour of users allow repetition but perhaps reduce validity. The difficulty and expense of any real system test suggests that much could be learnt from indirect tests, that is from simulation tests or from gedanken experiments.

The development of mechanized retrieval systems in the last two decades has effectively introduced a further categorization of tests, into those concerned with retrieval systems where the essential information characterization and search processes are done by human beings (though clerical operations may be automated), and those where these processes are to a greater or lesser extent taken over by the machine. The difference between these two types of system have implications for direct testing in particular.

The first four chapters in this section therefore consider testing, and especially experimentation, in the direct test contexts respectively represented by operational manual and mechanized systems, and by laboratory manual and mechanized systems. Lancaster discusses the problems of operational, essentially manual system studies, while Barraclough emphasizes some specific challenges presented by automatic systems, and particularly modern online systems. Keen considers the issues to be tackled in the design and conduct of laboratory tests of manual systems, while Oddy examines the special opportunities, but also dangers, of laboratory experiment with automatic retrieval systems.

To complement these treatments of direct testing, Heine examines

103

simulation testing in information retrieval, and Cooper concludes the section with a chapter putting the advantages of gedanken experimentation.

As in Part 1, there are many examples of individual tests intended to provide concrete illustrations of the problems encountered and methods adopted in the different types of test.