

Table 9.3
Collection Statistics

Number of documents		2504
Size of the descriptor dictionary in ARS		862
Average length of documents in the collection (in words)		94.6
Number of queries in the static collection		28
Number of queries in the dynamic collection		23
Average number of pertinent documents in the collection (per query)		15.4

The first experiment tested the feedback algorithm for a static collection of documents. The original search for each of the 28 queries was based on the query formulation constructed by the algorithm for the automatic construction of query formulations (see Chapter 7). The original user's requests consisted of marked sets of documents varying in size from four to six documents. Table 9.4 summarizes the results of this experiment.

The results in Table 9.4 were averaged for five different zones: queries with recall levels (1) less than 20%, (2) between 20% and 40%, (3) between 40% and 60%, (4) between 60% and 80%, (5) between 80% and 100%. The right side of Table 9.4 shows the averaged results for recall and precision levels after the feedback process. To demonstrate how the averaging was done for all zones, we chose one zone (20% to 40%) and tabulated all of the results in Table 9.5.

As the result of feedback all queries showed an improvement in recall levels, and 26 queries had an improvement in precision levels. The maximum number of iterations was seven (in two cases) and the minimum number was one (in four cases).

The second experiment was performed on the dynamic collection of documents. The collection was divided into three separate collections consisting of 850, 850, and 804 documents, respectively. There were 23 queries and the original search was performed on the first collection using query formulations constructed in the same way as done for the static collection. Table 9.6 gives the results of this experiment.

Table 9.4
The Results of the Experiment on the Static Collection

Recall	Precision	Final results after feedback	
		Original search	Precision
20	57.8	69.5	64.7
40	50.0	78.6	55.2
60	46.2	76.8	58.4
80	30.5	90.5	50.4
100	9.4	96.2	36.2

9. Control and Feedback in IR Systems

An illustration of the Averaging for the Recall and Precision Levels

Table 9.5

Number of relevant documents	Number of documents in the output	Feedback iterations						Initial search
		1	2	3	4	5	6	
11	3	5	27.2	37.5	-	-	-	44.4
19	6	4	31.5	60.0	1	0	94.7	62.0
15	4	3	26.6	57.1	4	-	66.6	58.8
12	4	4	33.3	50.0	2	0	66.6	66.6
13	5	4	38.4	55.5	3	1	69.2	53.3
21	7	5	33.3	58.3	1	0	85.7	60.0
14	4	6	28.5	40.0	2	1	71.4	47.6
17	6	8	35.2	42.8	4	1	94.1	51.6

An illustration of the Averaging for the Recall and Precision Levels

Table 9.5