Table 9.6
The Results of the Experiment for the Dynamic Collection

Recall level (%)	Precision (%)		
	Original search	First iteration	Second iteration
20	61.4	72.3	76.2
40	53.8	58.1	63.1
60	49.3	55.7	59.0
80	26.5	32.2	38.4
100	5.6	14.8	17.3

The experiments just described showed that the suggested feedback algorithms increase both recall and precision levels of a search. In other words, as the result of the feedback process, the larger part of POIN (area A in Figures 9.1 and 9.2) is covered, with the reduction in noise represented by the area outside of area A.

The algorithms described allow the system to not only adapt to the user's POIN but also to completely automate this process. However, as was mentioned earlier, this type of feedback does not solve the problem of optimal search. For this it is necessary to develop a different mechanism, which we discuss next.

9.7 Feedback for Search Optimization

In describing the function of IR systems (see Chapter 7), we discussed the need for an optimal search to satisfy the user's POIN. In the same chapter we considered the structure of an IR system realizing this function and identified the element responsible for this function. In this section we consider the structure of such an element, which will allow us not only to realize the process of optimization but also to fully automate this process. In other words, we will describe the algorithms used to select the system's best state for the static and dynamic collection of documents (Frants, Shapiro, & Voiskunskii, 1993; 1996). But before describing the algorithms, we first discuss the problem of search optimization.

Success or failure in a fully automated search depends on the quality of search algorithms as well the quality of the IRL used. It is in these two directions that developers of IR systems concentrate their attention. They try to improve the quality of the search by, on the one hand, improving the grammar and vocabulary of the IRL and, on the other hand, developing and improving the search algorithms. It should be noted that the development of IRLs proceeds in

the direction of bringing them closer to natural languages, and the development of search algorithms proceeds in the direction of establishing a more extensive consideration of the semantics of these languages. However, even when the information retrieval language is a natural language and there is an ideal search algorithm, failures in the search are possible. Here is why.

Because the goal of a search in an IR system is the information needed to satisfy the user's POIN, then it is quite obvious that the more precisely information about POIN is represented in the IRL (i.e., the more precisely the search request—and, in our case, the query formulation—represents the POIN), the more successful the search should be. However, obtaining a precise search request (query formulation) is difficult in principle. In fact, attribute 2 and, to a certain extent, attribute 3 of a POIN indicate that information about the POIN received by the system and the POIN itself generally differ from each other. Thus, even if we had an "ideal" method for constructing query formulations—i.e., even if we could absolutely precisely translate information about the POIN from natural language into the IRL—the query formulation obtained in this case would not ideally represent the POIN in the IRL. Moreover, the less the POIN and the search request correspond, the worse the search is, in spite of the "ideal" method of constructing query formulations. In creating IR systems, we cannot assume that in most cases users' requests will closely correspond to their POIN. Therefore, it is necessary to understand how the IR systems can function in this environment, what we can hope to achieve in satisfying users information needs, and how to realize these goals.

Today the efforts of the IR systems developers with respect to improving the quality of information service for users are concentrated in the direction of creating new, better methods and replacing the old methods with the new. In other words, the developers of the functioning IR systems use the best (from their point of view) methods (algorithms) for all processes existing in the system. However, as we showed earlier, even the creation of ideal methods does not solve the problem of a quality search. This by no means implies that better methods need not be created. On the contrary, their creation is necessary; but it is not sufficient. For the organization of a quality search, one must take into consideration attribute 2 of POIN.

Now we return to the function of an IR system. Recall that it was defined as follows: The function of a documentary information retrieval system is to fulfill an optimal (from the user's point of view) retrieval of information to satisfy this user's POIN with any information about the user's POIN given to the system. It should be pointed out that the requirement about *any information* about the POIN is a consideration of attribute 2 of POIN, which means that regardless of how well (or how badly) the request is stated, the user must obtain optimal output from the system. Recall that the concept "optimal" always means the best of all possibilities in the framework of a given system. Because "optimal" is the best of all possibilities, then consequently more than one possibility