tions. The values of r^2/n are computed for the outputs obtained from all subrequests of the best query formulation, and these values are stored for further analysis. After the search in the new collection of documents using the corrected query formulation (the first iteration) and the user's evaluation of the output, the values of r^2/n are computed for the outputs corresponding to the corrected query formulation and all its subrequests. Then the values of r^2/n for the corrected query formulation and for the previous query formulation are compared. If the quality of the search improved, then the query formulation is corrected and the next iteration is performed. If, on the other hand, the quality was lower, then all individual subrequests of the used query formulation are analyzed. Then the subrequests with a lower search quality than the previous query formulation are selected and divided into two sets. The first set contains all subrequests that were included in the previous query formulation, and the second set contains subrequests that were added after correction. If the first set is larger than the second, then this might indicate (especially, when the second set is empty) that the user's POIN changed and this change was correctly "observed" by the algorithm. Then the query formulation is corrected (for example, by removing the worst subrequests), a new search is performed, and so forth. However, if the second set is larger, then this could mean that the chosen algorithm was not the most appropriate. Hence, a new correction is performed for the previous algorithm and the algorithm with the next highest value of $\sqrt{R \cdot P}$. Then the search is conducted using two (corrected) query formulations, the comparison of the search quality is performed for both query formulations and their subrequests, and so on. If two algorithms have the same highest values of $\sqrt{R \cdot P}$, then the next iteration is performed on the basis of both algorithms. When eventually only one algorithm will give the best evaluation results, we will proceed with this algorithm as previously described.

It is clear that the proposed approach is only one way to realize internal control and, as before, it is important to realize the existence of different approaches to automatic internal control for both the static and the dynamic collection of documents. It is important to point out that the internal control, which becomes feasible with the introduction of the new characteristic $\sqrt{R \cdot P}$ of search quality, is an important part in the process of selective feedback.

9.12
Conclusion

It is difficult to overestimate the importance of feedback for the successful functioning of an IR system. Moreover, because the user is taking part in this process, an automatization of feedback becomes especially important. Such an automatization is a simple and convenient interface between the system and the

user because there are no decisions (methodological or strategical) that the user has to make. The user's only responsibility is to evaluate the documents in the output for their pertinence and to provide this information to the system in a natural language (for example, by responding yes or no).

It should be emphasized again that it is possible to organize an optimal search for an individual user in any system, not only the ones that use Boolean search. In this connection, we will indicate some additional interesting possibilities. We mentioned earlier that for performing an optimal search it is more expedient to use a set of different methods for constructing query formulations. But today it is also realistic to use more than one method for indexing documents. In such a case it is not necessary to use a special method of indexing for each user in order to create a personalized collection of document profiles to be searched. The system may contain several versions of indexed documents, which will be used in a search process as follows.

Assume that in some IR system the documents are indexed using two different methods. This could be done in cases where there are two different descriptor dictionaries or when one descriptor dictionary is used together with free text search, for example. Then the system will contain two separate search collections representing the same collection of documents. In such a system the choice of the optimal system's state for an individual user means a choice of the search collection that gives better search results. Clearly, in some cases the best could be considered a search performed on both collections. To realize this approach, it is possible to use algorithms that are similar to the ones described earlier. In this case the initial search for each search request will be performed on both search collections and the combined output will be given to the user for evaluation. Then using the suggested criterion of selection (r^2/n) it is possible to obtain the best system's state.

In the preceding example, only two search collections were mentioned in order to keep our discussion simple. It is clearly possible to use many search collections corresponding to the same collection of documents and moreover to use them in combination with several methods for indexing search requests. This will increase the number of choices and this, in turn, will improve the quality of service provided to the users.

In this chapter we considered the construction (algorithms) of the last elements of an IR system (see Figure 4.8 in Chapter 4), that is, the elements realizing control in IR systems. These elements realize the type of control that leads to the optimal service for an individual user (thereby fulfilling the system's function described in Chapter 4) and also allow the system to adapt to the user's POIN, specifically, to take into account attributes 2, 3, and 4 of POIN. Therefore, now we have a complex of interconnected algorithms capable of realizing a fully automated IR system, thereby fulfilling the system's function to the maximum. Before creating such a system, however, another point has to be considered. Recall that one of the very important steps in creating any system is evalu-