of a language in which the search is carried out. *Any search can be carried out only with the help of a language*, natural or artificial (for example, the language of mathematics, the language of signals, the language of form, or of color). The set of objects among which the search is carried out must be represented in the IN language of the search. The need for which the search is carried out also must be formulated in this language, then comparison must be carried out (by a person or a mechanism specially created for this purpose) of the set of objects with the formulated needs. Of course, the rules for comparison are determined by the the language in which the objects of the search and the need are given. If as a result of the comparison some objects correspond to the chosen criterion, then these objects are considered to be found. The following example illustrates a search process.

In searching for specific edible roots, our distant ancestor determined their location, for example, by the form or color of the stem or of the leaves. Of course, our ancestor must have known the language of forms and the colors of plants. By comparing a plant (represented by an object of specific form and color) with the theoretical image of the desired plant (represented in the same language), our ancestor determined the degree of their similarity and made a choice. On the other hand, suppose that during a search for food, our ancestor tried to determine the location of an animal by its characteristic cry (quacking). In other words, the first hunter was also familiar with a specific language of sounds. However, it is obvious that the hunter did not try to formulate this need in the search for roots in the language of sounds and in the search for animals in the language of colors and forms. Our ancestors clearly understood which language was needed for which search.

Schematically any search can be represented in the form of Figure 3.1 (adapted from Salton & McGill, 1983). It is obvious that this scheme is useful in

Process of determining the similarity of the need for a search with the object of the search base
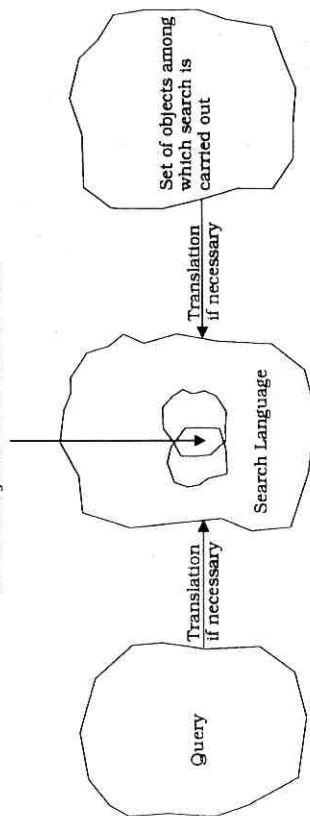
Figure 3.1
Block-diagram of the search process. Source: Adapted from G. Salton and M. J. McGill. *Introduction to Modern Information Retrieval* (New York: McGraw Hill, 1983). p. 11.

the search for edible roots, in the search for new planets in the universe, and in the search for a wife or a husband. It is useful also in the searches that interest us, particularly for *information retrieval*.

Note that the addition of the word "information" to the word "retrieval" relates to the goal of the search. In fact, because the product of satisfying an IN is information, *the goal of information retrieval is to find the information required to satisfy the IN*.

In our analysis of the concept "document," we noted that within the framework of the problems considered in the present book, in any discussions concerning information retrieval (except for those specially stipulated) we mean information contained in documents.

The notion of information retrieval has been known for a long time. Many scrolls in the Alexandrian library had titles, that is, a special element of text intended for search. By the title of a book or article, the reader can judge in the majority of cases whether this book or article will interest him or her and whether it is worth getting acquainted with in more detail. In the same library in about 250 B.C., Callimachus put together a catalog of the library's holdings— a volume of 120 scrolls. Nevertheless, in spite of ancient historical roots, even today the concept of information retrieval often causes discussion. As a rule, the discussions concentrate around the following questions: "What does it mean 'to seek information'?" "What in fact are we seeking?" In order to answer these questions we will begin considering information retrieval from the general notion of searching (see Figure 3.1).

First we recall that in our discussion the physical objects involved in information retrieval are written documents. Further, since retrieval can be carried out only in some language (natural or artificial), then the language in which the documents are written, or some other language, can be used as the search language. It is obvious that in the latter case a translation of the content of the documents into the same language in which the search is carried out is necessary. In information science, *information retrieval language (IRL)* is a language in which information during the search process is expressed and which is the basis for selection criteria. Of course, for a search process the need for information must be formulated in the same language. In this connection, assuming that any information need, the need for retrieval is carried out only because the user has an information need, the need for retrieval is the result of the user recognizing the existing IN. Moreover, users who carry out the search on their own do not have to express (for example, to pronounce aloud) this recognition for outside observers. Obviously, the actual process of information retrieval involves comparing objects of the search with the need for them.

We will consider the case in which the user chooses the necessary information, that is, the case when the user carries out the comparison without outside help. How and what does the user compare in the search that is being carried out? The user reads a document, somehow interprets its content, and evaluates (compares) this content with respect to the uncertainty which is the