level of development can be judged by the extent of the libraries and the high level of classification of sciences of that time, as well as by some indirect evidence contained in the literary memorials of ancient times.

It should be noted that the development of the library-bibliography classifications was in many ways moving in parallel to the development of science. At different times, various researchers have significantly influenced this development. For example, the works of K. Gesner, a Swiss librarian (1548), F. Bacon, an English philosopher (near the end of the 16th century), G. Leibnitz (the 17th century), and others may be noted. The fastest progress in the development of library classifications was made at the end of the 19th century and in the first half of the 20th century. These classifications have influenced both the arrangement of documents in a collection and the catalogs that reflect this arrangement. But the catalogs using these classifications constituted only auxiliary (albeit important) means of retrieval that, while playing an important role in improving information retrieval, still involved using the natural language for which retrieval was performed by humans. As we indicated in Chapter 3, the current information crisis has presented problems that the traditional classification methods could not solve. This is exactly why we will not discuss in great detail the classification languages that are typically used in the traditional library methods but will move, instead, to the consideration of languages used in the IR system.

5.3 IRL for the IR System

In Chapter 4, we noted that the retrieval of information in the IR system, that is, direct operations of comparison and selection of information, is performed using a computer without the participation of people. It is understood that the absence of humans during a retrieval means the impossibility of using a natural language for the retrieval. Thus, the *creation of an artificial language for the IR system is a necessity*—at least until such time as when an artificial intellect that is able to master the semantic component of a natural language is created. With this in mind, note that the creation of such an intellect does not look very promising today. But what should an artificial IRL be like? The following exposition is an attempt to answer this question.

Based on the fact that during the process of retrieving information, the IRL must represent the objects of the retrieval, as well as the search requests (see Chapter 2), we will begin by defining *what* should be understood by such representation in an artificial IRL, that is, what must be represented and how it may be represented. Apparently, for the purpose of performing a quick retrieval, IRL must, to the fullest extent, take into account the specifics of the objects to be retrieved and the requests for the retrieval. (From this point, when we say IRL

we will mean only artificial IRL intended for use in the IR system.) It is clear that the objects of a retrieval in the IR system are the documents in the collection. Let's recall that a *document* refers to a material carrier with attached information. Specifically, the IR system deals with documents whose material carriers contain a text entered with the help of the written component of a natural language. (It should be noted that out of more than 3000 natural languages currently used by people, fewer than 10% have written components.) Obviously then, in order to perform a retrieval, the IRL must represent not the material carriers but only the information that is attached to them. This representation must be "something" that would allow the required information to be found, and this "something" has to be different for different types of information.

But what differentiates one type of information from another? First is the difference in how each reflects the fragments of reality. Based on the fact that the information in written documents is encoded in the form of text communications using a natural language, after the reading of the texts, these differences are perceived by humans as different meanings of those communications. Because only the meanings of communications present an interest for the users, it is precisely the meanings that must be found in a retrieval. Thus, in submitting the documents to the IRL, we are, in a sense, attempting to somehow present (encode in the IRL) the meaning of the communications attached to these documents. It should also be remembered that the representation of meaning in the IRL must be convenient for a retrieval performed on a computer and not for some other goals. For example, if the available representation seems sufficient for a high-quality search, we should not be concerned with how the quality of the meaning will be perceived by the person reading such representation.

We noted earlier that the IR system for the IRL must represent not only the documents, but also search requests. It is apparent that in formulating search requests humans also encode some meaning in the communication (request) and basically rely on the fact that the documents containing communications with similar meaning will be found in a search process. In this case, the meaning contained in the request must somehow be represented in the IRL.

Thus, we have shown that the meaning of the documents from the collection of documents, as well as the meaning of requests entered into the system by a user, must be represented in the IRL. These representations are then compared using some formal criterion and the necessary information is selected. We would like to point out that in this process, the computer is not operating with meanings but with the representations of meanings. Of course, they are not the same. After all, it does not surprise us, for example, that an elevator button represents a floor, a social security number represents each person legally residing in the United States, and so forth. Now, after understanding *what* must be represented in the IRL, let's consider *how* it may be done.

Many linguists believe that artificial languages are created and exist only on the basis of natural languages. The existing practice supports this opinion. In