

quality of service to the user, and it points to the necessary existence within the structure of the system of an element fulfilling this requirement. Moreover, because optimization is one of the tasks of control, it is necessary to consider a documentary IR system as a controllable system (Frants & Shapiro, 1991). We recall that in a controllable system an indispensable condition of any optimization is the existence of feedback. Consequently, the requirement of carrying out optimal information retrieval assumes the necessary existence of the element of feedback in the structure of a documentary IR system. The necessity of the existence of feedback in the system also follows from another important factor implied from the definition of the function of a documentary IR system. This factor is the following condition: it is not the information need in general that must be optimally satisfied, but specifically the POIN of the user. In other words, in the realization of the function of a documentary IR system it is necessary to consider the properties of POIN.

In analyzing the influence of types of IN on the types of IR systems, it was pointed out that consideration of attributes 2 and 4 of POIN assumes existence in an IR system of a mechanism that decreases the discrepancy between the user's existing POIN and its representation in the system. It was mentioned that feedback must be used for realization of this mechanism. However, we did not dwell on why feedback is necessary in this case. Therefore we will discuss this in detail.

We will consider a case in which users' search requests enter the system. This is not only the most frequently encountered but also the most expedient situation, which will be considered in detail in Chapter 7. We recall that attribute 2 of POIN tells us that, as a rule, the system receives information that does not reflect POIN precisely; that is, the information contained in a search request does not correspond to POIN exactly. However, now, after our consideration of BIQ in Figure 4.3, we can say that there is also another source that causes the system to use imprecise information about POIN in its retrieval process. The issue is that inside the system the user's search request is translated from a natural language into an IRL (if the IRL is distinct from the natural language). Today in documentary IR systems, different methods of indexing search requests are used (or, as it is often said, different methods of constructing query formulations), but all of the methods are far from the ideal. Given the same search request each of the methods may (and most often does) result in a different query formulation, which will represent the actual POIN of the user differently in the system. Because retrieval of information is carried out on the basis of the constructed query formulation in the system, the output, generally speaking, corresponds to some hypothetical POIN represented by the query formulation and not to the actual POIN of the user (Frants & Shapiro, 1993).

We will illustrate what we have just said with the help of Venn diagrams. In Figure 4.4(a), region A represents the actual POIN of a user. (By actual we mean in fact a POIN that must be satisfied as a result of a search.) In Figure 4.4(b), a

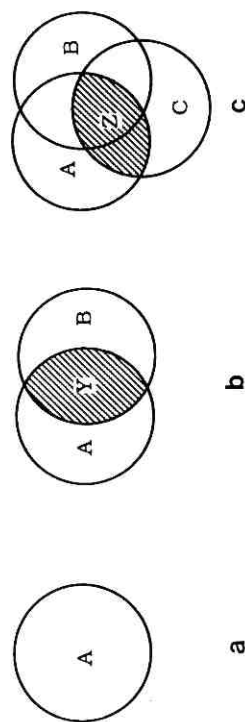


Figure 4.4

Example of the discrepancy between actual POIN and its representation in the user's search request and IR system.

new region B represents the expressed POIN given in the user's search request. In order words, by formulating a search request, the user "shows" the system the need that in some way corresponds to A; that is, the user thinks that he expresses A, although he actually expresses B. Area Y (the intersection of A and B) shows which part of the actual POIN was given to the system. In Figure 4.4(c), a new region C is introduced, which represents the satisfied POIN, that is, the POIN that will be satisfied by the system's output. Area Z (the intersection of A and C) shows which part of the actual POIN was satisfied.

But the system's goal is to satisfy the actual POIN (represented by A). Hence, after the output satisfying region C is obtained, it is necessary to have some mechanism to determine the existence and the level of discrepancy between A and C. Afterward, this mechanism will generate controlling signals to change the system's state so that the system's next output will correspond more closely to the actual POIN. But the mechanism minimizing the difference between the action's goal and its result is feedback in its general form (see the definition for feedback in Chapter 1). Hence, to generate controlling signals we need to use the feedback mechanism. Thus, the necessity of control with the help of feedback follows also from the requirement to closely match regions A and C.

The character of control in this case is different from optimization, which is a special case of adaptation. But what kind of control is it? We will consider this question in more detail.

It is known that in control systems, four basic types of control problems are addressed: stabilization, program control, monitoring, and adaptation. Hence, in discussing control systems the following terminology is used: "systems of stabilization," "systems of program control," "servo systems," and "adaptive systems." The most complicated type of control is adaptation, and that is exactly the type of control that is characteristic for a documentary IR system. Why should an IR system be considered an adaptive system?