can be determined by the angle between them. This is also confirmed by the CSCs produced in the cases discussed.

Now let us consider one more model, which allows production of CSCs similar to characteristic I_6 (that is, we will demonstrate one other way to construct CSCs similar to I_6, as mentioned earlier). An interest in such models is explained by the rather large number of complex search characteristics, similar to characteristic I_6, proposed by other authors. Among these is the CSC

$$I_9 = 1 - \frac{1}{\frac{1}{R} + \frac{1}{P} - 1},$$

which was proposed by Heine (1973); the CSC

$$I_{10} = 1 - \frac{1}{\frac{2}{R} + \frac{2}{P} - 3},$$

which was proposed by Vickery and documented by Cleverdon, Mills, & Keen (1966); and the CSC

$$I_{11} = 1 - \frac{2}{\frac{1}{R} + \frac{1}{P}},$$

which was proposed by van Rijsbergen (1979). It is hardly expedient to consider all such models. Therefore, we will focus on the model developed by van Rijsbergen (1979), which is attractive because of its simplicity.

Now, by A let us denote the set of documents found during the search, and by B let us denote the set of pertinent documents in the search collection. The relationship between these sets represents the results of the search as illustrated in Figure 10.8

In this figure, the unshaded area represents the set of obtained pertinent documents during search, the shaded area A represents the set of obtained nonpertinent documents, and the shaded area B represents the set of pertinent documents not found during search. Clearly, the higher "degree of coincidence" of sets A and B, the better the search result. For example, in the case of a full coincidence of these sets, we get an "ideal" output. It is the tools used to determine the "degree of coincidence" of sets that are proposed for the evaluation of search results within the discussed model.

The following criterion seems to be a "natural" tool to determine the "degree of coincidence" of sets M and L:

$$\frac{|M \cap L|}{|M \cup L|},$$

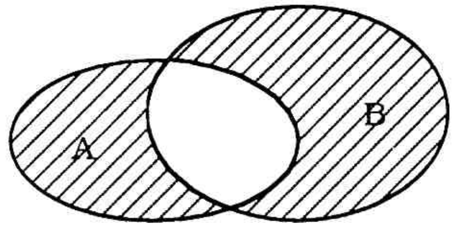

Figure 10.8
Relationship between sets A and B.

where M and L are finite sets, $|M \cap L|$ is the number of elements in the intersection of sets M and L, and $|M \cup L|$ is the number of elements in set union of M and L. Applying this criterion to sets A and B, keeping in mind that $|A \cap B| = r$ and $|A \cup B| = N + C - r$, and assuming that $r \neq 0$, we get the following:

$$\frac{|A \cap B|}{|A \cup B|} = \frac{r}{C + N - r} = \frac{1}{\frac{C}{r} + \frac{N}{r} - 1} = \frac{1}{\frac{1}{R} + \frac{1}{P} - 1}.$$

In other words, by application of the suggested criterion we arrive exactly at the complex search characteristic I_6. At the same time, van Rijsbergen used the following criterion as a tool to determine the "degree of coincidence," or rather the "degree of noncoincidence" of sets:

$$\frac{|M \cup L - M \cap L|}{|M| + |L|},$$

which, when applied to sets A and B, gives complex search characteristic I_{11}. This is shown as follows, assuming again that $r \neq 0$:

$$\frac{|A \cup B - A \cap B|}{|A| + |B|} = \frac{N + C - 2r}{N + C}$$
$$= 1 - \frac{2r}{C + N}$$
$$= 1 - \frac{2}{\frac{C}{r} + \frac{N}{r}} = 1 - \frac{2}{\frac{1}{R} + \frac{1}{P}}.$$

It should be kept in mind that it follows from the construction method of the described CSC that the higher the search results, the lower the values that are obtained by the characteristic. This completes our discussion of the model suggested in the work of van Rijsbergen (1979).