Then,

$$I'_{12} - I_{12} = \frac{r'_{i1}d'_{i1} - l'_{i1}b'_{i1}}{\sqrt{(r'_{i1}+l'_{i1})(r'_{i1}+b'_{i1}+d'_{i1})(l'_{i1}+d'_{i1})}}$$

$$- \frac{r'_{i2} \cdot d'_{i2} - l'_{i2} \cdot b'_{i2}}{\sqrt{(r'_{i2}+l'_{i2})(r'_{i2}+b'_{i2}+d'_{i2})(l'_{i2}+d'_{i2})}}$$

$$= \frac{8 \cdot 1030 - 12 \cdot 2}{\sqrt{20 \cdot 10 \cdot 1032 \cdot 1042}} - \frac{6 \cdot 1036 - 6 \cdot 4}{\sqrt{12 \cdot 10 \cdot 1040 \cdot 1042}}$$

$$= \frac{8}{\sqrt{10 \cdot 1042 \cdot 480}}\left(\frac{1030-3}{\sqrt{20 \cdot 1032}} - \frac{777-3}{\sqrt{12 \cdot 1040}}\right)$$

$$= \frac{8}{\sqrt{10 \cdot 1042 \cdot 480}}\left(\frac{1027}{\sqrt{43}} - \frac{774}{\sqrt{26}}\right)$$

$$= \frac{8}{\sqrt{10 \cdot 1042 \cdot 480}} \times \frac{\sqrt{1027 \cdot 1027 \cdot 26} - \sqrt{774 \cdot 774 \cdot 43}}{\sqrt{43}\cdot\sqrt{26}}$$

$$= \frac{8}{\sqrt{10 \cdot 1042 \cdot 480}} \cdot \frac{\sqrt{27422954} - \sqrt{25760268}}{\sqrt{43}\cdot\sqrt{26}} > 0.$$

Thus, we see that in Situation 1 and in Situation 2 the difference $I'_{12} - I_{12}$ has opposite signs, and the only parameter distinguishing two situations, as previously, is the number of pertinent documents in the search collection, that is, only in values of C. Hence, the sign of difference $I'_{12} - I_{12}$ depends on the value of C. This means that complex search characteristic I_{12} does not have order preservation property.

Thus, we have determined whether certain CSCs have the order preservation property. In conclusion it should be pointed out that, as was stated before, characteristics I_2 and I_{12}, in spite of the proximity of the values under certain conditions, do differ significantly, as became evident during the consideration of the order preservation property as applied to these CSCs.

10.9

Fuzzy Scales of Pertinence

As we discussed in the first section of this chapter, during the pertinence analysis of a document by a user, situations may arise when a user would find it so difficult to judge whether the analyzed document is pertinent that he or she would consider it helpful for evaluation of document search results to take into account the nonpertinence degree of nonpertinent documents. In this case,

some of the possible choices would be almost pertinent, possibly pertinent, almost nonpertinent, and so on. Naturally, the binary scale of pertinence is out of the question here. Rather, the fuzzy scale is used, which, in addition to the pertinent and nonpertinent values, includes "almost pertinent" and "almost nonpertinent," among other values. (All these values can also be expressed numerically.)

In discussing the possibility of using fuzzy scales of pertinence, it should be pointed out that taking into account the nonpertinence degree of nonpertinent documents in the process of evaluating the quality of a produced output is associated with a somewhat different understanding of the functional efficiency of document search, because in the previous discussion of the quality of the output and, hence, functional efficiency, we did not consider the nonpertinence degree of nonpertinent documents. It follows that it is necessary in this situation to specify the content criteria of a functional efficiency evaluation. (Recall that in the first section of this chapter we mentioned the scale of pertinence employed by the user in analyzing how a document affects the development of concrete content criteria.) Regretfully, neither informational practice nor an analysis of relevant publications help in establishing these criteria. What's more, in our opinion, researchers do not have a clear picture of these criteria, even on a intuitive level. (This is the result of the fact that, in practice, fuzzy scales of pertinence, as a rule, are not used.) We are also not prepared to define a set of content criteria of functional efficiency evaluation that would provide an adequate understanding of the position from which functional efficiency could be evaluated in the given situation. We note that to solve this problem, one must have a clear idea of what should be taken into consideration in an output quality evaluation: the nonpertinence degree of both found and not found nonpertinent documents or, say, only that of found nonpertinent documents, and so forth. We, too, have no clear understanding of how this should be taken into consideration. There is a point of view that one should not take into account the nonpertinence degree of not found nonpertinent documents when evaluating output quality. The attractiveness of this alternative is explained by the following way. All nonpertinent documents not found during the search are not included in the output, in our opinion, with equal legitimacy, irrespective of their nonpertinence degree. However, other points of view could also have persuasive arguments. Because at present nobody has exhaustively analyzed this problem, let us choose the most convenient point of view; namely, let us assume that in output quality evaluations one must take into consideration the nonpertinence degree of both found and not found nonpertinent documents.

Keeping this point of view in mind, let us consider only one content criterion of the functional efficiency evaluation (we do not claim that this criterion could be used without any reservations). In the subsequent discussion, we will assume that in analyzing documents for their pertinence a fuzzy scale of pertinence is used with m possible choices: $f_1, f_2, \ldots, f_m$, where $f_1 = 1$ (a document is pertinent), $f_m = 0$ (a document is nonpertinent), and f_i—for $2 \leq i \leq m - 1$—