Thus, two parts Δ_1 and Δ_2 of K are regarded as duplicates if they contain subsets $A \subset \Delta_1$ and $B \subset \Delta_2$ such that the cards from A and B were originally among the copies of the same sufficiently small, connected, piece of the original deck. Note that Δ_1 and Δ_2 may contain no identical cards at all, since it is possible that $A \cap B = \varnothing$. However, in shuffling incompletely, there must be copies of Δ, distributed in K with certain cards from A and B not far from each other, which means, in the case where Δ_1 and Δ_2 contain fragments arising from the common inverse image of Δ, that the probability increases of two cards from Δ_1 and Δ_2 respectively, being close somewhere in K. This fact can be used for making more precise the concept of 'similarity' of pieces in K and introducing a measure relation for them on the basis of the quantity of such card meetings.

We now perform the detailed argument for *long chronological lists*. Let there be a list K which may contain errors, omissions, and/or duplicates. We denote by Y an unknown original list on which X is based. Thus, Y is an imaginary list containing complete data of a certain sort (say, about the rulers' names) for a long historical time interval I_Y. Let I_Y be described by a number of chroniclers, each making his own short list Z for the contemporary events. Denote by $\{Z_i\}$ the set of these short lists, forming a certain covering of Y, assumed to be sufficiently dense (with large multiplicity), and containing somehow dispersed and, possibly, erroneous pieces, each of Z_i mentioning not all the rulers' names or not every one of the persons; besides errors and gaps could occur in rewriting and compiling which, for simplicity, we will assume intrinsic to Z_i from the beginning.

In making the chronology in its contemporary form, the result was a new arrangement of Z_i, and the known list X was obtained. Consider two intervals Δ_1 and Δ_2 in X. Let us try to determine whether or not there is a pair Z_i, Z_j in X, which would be related to one period in Y, and glued to Δ_1 and Δ_2, respectively. As in the example with cards, we conclude that if there is such a pair, then the probability increases that the names from Δ_1 and Δ_2 will be proximal somewhere in X, on account of a third, 'gluing' chronicle Z_l. We will use the term 'glue' to describe the procedure of identification of similar pieces in different fragments of the list.

12 Relation of a pair of names

For the time being we neglect the partition of a list into chapters. In contrast to the problem of determining the shift values, to construct the relation matrix the time scale was not used in the list. After constructing the matrix, we again make use of the analysis of the results.

To make the concepts of 'piece of a list' and 'proximity in a list' precise, we introduce the following definitions.

Definition 4. We call the set $\Delta_i(k) = \{a_{i-k}, \ldots, a_{i+k}\}$, $k < i < N-k$, the determining neighbourhood of radius k for the ith entry a_i of the list $X = \{a_i, \ldots, a_N\}$. We also call $2k+1$ the length of the determining neighbourhood, and do not introduce the concept for the extreme terms. We denote $\Delta_i(k)$ simply Δ_i, and sometimes omit the term 'determining'.

Definition 5. We call the number $l_0(u_i, u_j)$ of pairs (a_s, a_r), $u(a_s) = u_i$, $u(a_r) = u_j$ of noncoincident entries of the list X, so that $|s-r| < p$, the nonnormed relation. We also call the natural number p the length of the relating neighbourhood.

Parameters k and p were chosen in accordance with the list. Note that the general form of the relation matrix was invariable for all the values of k and p, $1 \leqslant k \leqslant 7$, $3 \leqslant p \leqslant 17$ considered, in all the above examples, so that this choice did not influence the result itself