following definitions.

Definition 4. We call the set

$$\Delta_i(k) = \{a_{i-k}, \ldots, a_{i+k}\}, \qquad k < i < N - k,$$

the *determining neighbourhood* of radius k for the ith entry a_i of the list $X = \{a_1, \ldots, a_N\}$. We also call $2k+1$ the *length of the determining neighbourhood* and do not introduce the concept for the extreme terms. We denote $\Delta_i(k)$ or 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 non-coincident entries of the list X, such that $|s - r| < p$, the *non-normed relation* of two names u_i and u_j. 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 considered values of k and p, $1 \leqslant k \leqslant 7$, $3 \leqslant p \leqslant 17$, in all the above examples, so that this choice did not influence the result itself (decomposition of the list into a duplicate system), but only its precision.

The non-normed relation $l_0(u_i, u_j)$ is inconvenient, because it does not take into account sharp differences in the multiplicities of the names from I, which are characteristic in the examples in question. Meanwhile, a pair of frequent names should naturally be at a close distance in X more often than a pair of rarer ones. To eliminate the influence of the multiplicity of names on their relation, we introduce the following definition.

Definition 6. Let two names $u_i, u_j \in I$ be in a list X with multiplicities k_i and k_j, respectively. We call the number

$$l(u_i, u_j) = \begin{cases} l_0(u_i, u_j)/(2k_i \times k_j) & \text{for} \quad i \neq j, \\ l_0(u_i, u_j)/(k_i(k_i - 1)) & \text{for} \quad i = j, \; k_i > 1 \end{cases}$$

the (*normed*) *relation* of a pair of the names u_i and u_j.

By definition, we put $l(a_r, a_s) = l(u(a_r), u(a_s))$ for $a_r, a_s \in X$. We chose the norming procedure in Definition 6, so that, assuming that for the given name set $I = \{u_1, \ldots, u_m\}$ with multiplicities $k_1, k_2, \ldots, k_m$, all permutations in the correct list X may be equally probable, (in other words the names in the chronologically correct list may be distributed at random, and the knowledge of only the name set with multiplicities does not supply any information regarding the particulars of their position in the list), and the relation of two names in X may be a random variable with mean not depending on the choice of a name pair. This (general) mean will be called mean with respect to the permutations in contrast with the empirical mean with respect to the matrix. This assumption is confirmed indirectly by the coincidence (in the correct lists) of the theoretically general mean α calculated by formula (3), with the empirical mean with respect to the matrix, whereas for the lists with duplicates, as had to be expected, the mean relation with the matrix is slightly greater than α. Note that the said assumption does not influence the qualitative form of the results. In particular, the basic features of the essential relation matrix are also preserved in using the non-normed values of the relation.