Bernoulli test sequence, and regard as a success in a trial if u_j gets into a relating neighbourhood along with one of the copies of u_i already placed. Then $l_0(u_i, u_j)$ equals the number of successes, and the probability of a success in one trial is proportional to k_i, whereas the number of trials equals k_j; therefore the mean number of successes is proportional to $k_i k_j$. The case $i=j$ is considered similarly. Thus, the mean $Ml(u_i, u_j)$ does not depend on the pair (u_i, u_j) except when $u_i = u_j$ is a unique name in the list.

13 Relation of two neighbourhoods in the list

In accordance with the assumptions of § 11 regarding the duplicate appearance mechanism, we introduce a relation measure for two determining neighbourhoods $\Delta_r(k)$ and $\Delta_s(k)$, $k<r$, $s<N-k$, in the list X, namely,

$$L_0(\Delta_r(k), \Delta_s(k)) = \frac{c}{(2k+1)^2} \sum_{i=r-k}^{r+k} \sum_{j=s-k}^{s+k} l(a_i, a_j), \tag{5}$$

where c is a certain constant which we will choose as convenient.

Definition 7. The number $L_0(\Delta_r(k), \Delta_s(k))$ is called the relation of two neighbourhoods $\Delta_r(k)$ and $\Delta_s(k)$ in the list X.

If X contains no duplicates, and the assumptions of § 12 are valid, then as seen from (5), the mean value of the relation $L_0(\Delta_r, \Delta_s)$ does not depend on Δ_r and Δ_s, and equals $c\alpha$, where c and α were defined in (3) and (5). Here we imply the mean with respect to placing with respect to r and s. Let us see how the value of the relation of neighbourhoods Δ_r and Δ_s is altered by their common duplicates. The following definition is introduced to make precise by how much the relation increases because of one of their common 'complete' duplicates in X, that is such a connected piece in X, which is not longer than the relating neighbourhood, containing the copy of each name from Δ_r and Δ_s (taking the multiplicities into account). If we represent Δ_r and Δ_s as two mutually completing rare chronicles having a common inverse image from Y, then the complete duplicate is the complete chronicle combining the names from these two.

Definition 8. We call the number

$$E_0(\Delta_r(k), \Delta_s(k)) = \frac{c}{(2k+1)^2} \sum_{i=r-k}^{r+k} \sum_{j=s-k}^{s+k} \frac{1-\delta_{ij}}{c(a_i, a_j)},$$

where $\delta_{ij} = 1$ if $i=j$ and 0 otherwise, and $c(a_i, a_j)$ was defined in (4), c the multiplier from (5), the proper relation unit for two determining neighbourhoods.

Let X contain duplicates. We call two determining neighbourhoods *independent* if they have no common duplicates and are nonintersecting in X. We call the remaining neighbourhood pairs *dependent*. We assume for simplicity that there are few duplicates, so that the relation between two independent neighbourhoods is similar to the correct list.

Consider the three following cases.

Case 1. The neighbourhoods Δ_r and Δ_s are *independent*. Then the mean value of their relation equals $c\alpha$.

Case 2. The neighbourhoods Δ_r and Δ_s *coincide*, Δ_r having no duplicates. The mean value of the relation in this case equals $c\alpha + E_0(\Delta_r, \Delta_s)$, and the neighbourhood is its complete duplicate.