In Dom Froger's book " La critique des textes et son automatisation" (1) the problem of orienting a given enchainment of texts, that is of finding the archetype has not been fully dealt with. This problem arises when all neighbour-connexions of the texts have been established (by a computer) using neither the value and "probability" of a variant nor the frequency of a constellation-type (2). Since better readings have not yet been differentiated from worse, the computer-output results in a diagram showing only the plain enchainment of the texts, i.e. steps of text-mutation, but not the direction of dependence between any of the texts. This enchainment has to be thought of as an arrangement with some special properties deriving from the later explained axioms. First however, the notation has to be set forth.

Notation and definitions

Let c_i be the i^{th} term of the collation-text C, where a term can be defined as a word, letter or morph, including the places between all neighbouring letters (where additions might occur). Then C is a sequence of all its n terms :

$$C = (c_i)_{i = 1, \dots , n},$$

and x_i being the reading of a text X for c_i, X too is a sequence of n terms :

$$X = (x_i)_{i = 1, \ldots, n} .$$

If c_i is defined as a single letter or place between two letters and x_i is not equal to c_i, x_i may be as well a sequence of words.

Let (X) be a reconstructed text, which has not been collated but which appears in the enchainment-diagram because it is represented by a list of variants,

[X] a lost text which cannot be reconstructed,
$\Omega := \{X\}$ the set of texts (supposed to represent the same work of a certain author) which have ever existed,
$(\Omega) := \{(X)\}$ the set of reconstructed texts,
$[\Omega] := \{[X]\}$ the set of unknown texts,
$\mathbf{\Omega} := \Omega - [\Omega]$ the set of the known texts, i.e. texts appearing in the enchainment-diagram as represented by a list of variants,
$m := |\mathbf{\Omega}|$ the number of these texts.

Further let $\rightarrow$ in $X \rightarrow Y$ be the relation of *direct dependence*, where Y has been copied directly from X, — in X — Y the relation of *direct enchainment*, where X has been directly copied from Y or Y from X, i.e. $X — Y :\Leftrightarrow X \rightarrow Y \vee X \leftarrow Y$ (3), and analogously $\rightarrow \ldots \rightarrow$ and — ... — the not (necessarily) direct dependence and enchainment (4).

Then *orientation* is just the exclusion of one of the two directions of dependence in $X \rightarrow Y \vee X \leftarrow Y$, based on the principle of accumulation of faults and the actual list of variants between X and Y.

Contiguous enchainments part from a common text, e.g. X — Y and Y — Z in X — Y — Z. A *chain* is a sequence of texts each of which is connected by a direct enchainment with the next one. The *complete enchainment of* Ω may be symbolized by (Ω, —) and the *stem*, i.e. the completely oriented enchainment by (Ω, $\rightarrow$).

Axioms

1. No text, directly or indirectly, can be copied from itself, i.e.
$$\neg (X \rightarrow \dots \rightarrow X),\ X = X.$$

2. No term (5) of a certain text X can have been directly copied from more than one term of another text, i.e.
$$\forall x_i \in X\ y_j \neq z_k \Rightarrow \quad (y_j \rightarrow x_i \leftarrow z_k).$$
If X is contaminated, it can be divided into r parts X_p each of which derives completely from a single source. In fact, it can be obtained that only these parts appear in the enchainment-diagram (6). We may assume therefore
$$\neg (Y \rightarrow X_p \leftarrow Z),\quad Y \neq Z,\ p = 1,..,r$$
or even
$$\neg (Y \rightarrow X \leftarrow Z),\quad Y \neq Z,$$
if we use different letters for the single parts of a contaminated text.

3. All copies derive from a common source (the *original* 0), i.e.
$$\exists 0 \epsilon \Omega\ \forall X \epsilon \Omega - \{0\} \quad 0 \rightarrow \cdots \rightarrow X$$

4. Older texts cannot have been copied from younger ones.

Results

From axiom 1 follows that

5. two texts X and Y do not depend on one another, i.e.
$$\neg (X \rightarrow Y \wedge X \leftarrow Y).$$

From axiom 2 follows that

6. there is only one original 0, i.e. $\exists$ in 3 is $\exists\,!$

From axioms 1 and 3 follows that

7. the original 0 does not depend on any text of the work considered, i.e.
$$\exists!\ 0 \in \Omega \quad \forall X \in \Omega \quad \neg (X \rightarrow 0).$$

From axioms 2 and 3 follows that

8. the number of direct enchainments in (Ω, —) is m - 1.

From axiom 3 follows that

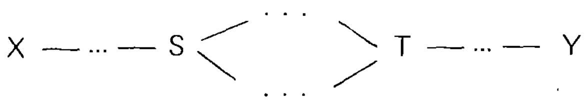
9. there is a longest chain 0 —— ... —— A, so that all known texts depend on A only (and its predecessors, 0 = A possible), i.e.
$$\exists!\ A \in \Omega \ \forall X \in \Omega - \{A\} \ (A \rightarrow \cdots \rightarrow X \wedge \neg \exists A' \in \Omega,$$
$$A \rightarrow \cdots \rightarrow A', \ \forall X \in \Omega - \{A'\} \ A' \rightarrow \cdots \rightarrow X);$$

this A is called *archetype.*

The archetype A may be one of the texts appearing in the enchainment-diagram, i.e. $A \epsilon \Omega$ (m possibilities), or it may not appear and be intermediate to two known texts X and Y appearing in a direct enchainment, i.e. $A \epsilon [\Omega]$ (i.e. $A = [A]$), $X — \cdots — [A] — \cdots — Y$, $X — Y$ subenchainment of $(\Omega, —)$ (m - 1 possibilities according to 8). Therefore a given enchainment $(\Omega, —)$ can be transformed by 2 m - 1 different ways of orientation into the same number of stems $(\Omega, \rightarrow)$.

From result 9 follows that

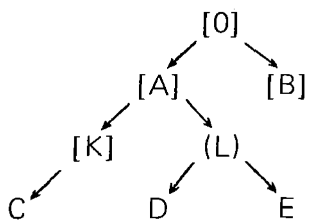
10. any two known texts X and Y are connected by a chain :
$X, Y \epsilon \Omega \Rightarrow \exists A \epsilon \Omega$, $P — Q$ subenchainment of $(\Omega, —)$
$A \rightarrow \cdots \rightarrow P \rightarrow \cdots \rightarrow X \quad \wedge \quad A \rightarrow \cdots \rightarrow Q \rightarrow \cdots \rightarrow Y$
($P = X$, $Q = Y$, $P = Q$, or $P = Q = A$ possible)
$\Rightarrow$ $P — Q$, $P — \cdots — X$, $Q — \cdots — Y$ subenchainments of $(\Omega, —)$
$\Rightarrow$ $X — \cdots — P — Q — \cdots — Y$ subenchainment of $(\Omega, —)$.
In the case of $P = Q$ or $P = Q = A$, there might be an equal number of identical direct enchainments to both sides of A or $P = Q$, i.e. those from A to R ($R \epsilon \Omega$) in $X — \cdots — R — \cdots — A — \cdots — R — \cdots — Y$ ($R = A$, $R = X$ or $R = Y$ possible, $R = R$); the omission of these identical direct enchainments produces the *shortest chain;* this is uniquely determined, otherwise there would be ramifications of the type



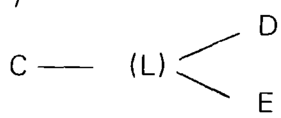
(X = S or T = Y possible), which in both directions of dependence would contradict axiom 2.

Example

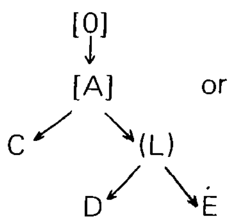
The real tradition (Ω, →)



The enchainment (Ω, —)



The stem (Ω, →)



or C = A or (L) = (A)

Transitivity of dependence

Given a dependence $X \rightarrow Y$ and an enchainment $Y — Z$ $(Y = Y)$, the enchainment $Y — Z$ is a dependence $Y \rightarrow Z$, otherwise $Y — Z$ would be $Y \leftarrow Z$, i.e. $X \rightarrow Y \leftarrow Z$ contrary to axiom 2, i.e.

11. $X \rightarrow Y — Z \Rightarrow X \rightarrow Y \rightarrow Z.$

By complete induction results

12. $X \rightarrow Y — \cdots — Z \Rightarrow X \rightarrow Y \rightarrow \cdots \rightarrow Z$

    In the condition of 11 and 12, $X \rightarrow Y$ may be the result of $X — Y \wedge X$ older than Y.

By 12 and axiom 2 results

13. $X \rightarrow \cdots \rightarrow Z \wedge X — \cdots — Y — \cdots — Z \Rightarrow$
    $X \rightarrow \cdots \rightarrow Y \rightarrow \cdots \rightarrow Z.$

By 13 the dependence of the intermediate texts of any shortest chain is induced by the relative dependence of the outermost texts of this chain.

By 12, the relative dependence of two neighbouring texts determines the direction of dependence not only of a single chain, but of all chains stemming from the dependent text and not including the independent of the two texts compared. For this reason, and because the list of variants between two neighbouring texts is smaller than that between one of these and a more distant one behind the other, it will be more convenient to establish the complete orientation of the stem (Ω, $\rightarrow$) by repeating step 12, i.e. always examining the dependence of two neighbouring texts.

Proceeding

Froger proposed to begin with the "upper" text, which would be the collation-text C, and to proceed always to one of the not yet oriented contiguous direct enchainments. It is possible however, that C is one of the outermost texts of one of the chains of maximal length l occurring in the enchainment (Ω, —) and that the archetype A is the other outermost text of this chain. In this case Froger will perform l-1 steps, and if at the ramifications he first examines those enchainments which do not determine the other, further steps might be added. To reduce this number, I propose the following systematical proceeding :

14. I) Pick an arbitrary chain of maximal length l within the unoriented part of (Ω, —).

II) a.- If l is even : examine the middle direct enchainment.
b.- If l is odd : examine one of the contiguous direct enchainments of the middle text (that with the larger list of variants will be more convenient, see below).

III) Exclude the dependent chains by 12.

IV) a.- If the remaining unoriented part of (Ω, —) contains more than one direct enchainment, go to (I).

b.- If there remains a single direct enchainment, the direction of dependence should be determined by 16 and 19.

Already the first step will have the same effect as the first l/2 (l even) or (l-1)/2 (l odd) steps of Froger (at least). After these Froger might have reached the middle of a chain of the new maximal length (within the unoriented part of (Ω, —)); then the next step, and thence its effect, in both methods would be identical, otherwise there would be a new disadvantage for Froger, etc.

Tests of independence and dependence

Assuming the principle that the original 0 is correct and that copyists add faults, Froger held that a single fault in a list of variants sufficiently proves that the texts with the false reading have been copied from another text and may be eliminated as dependent; that is correct in Ω, but misleading if the original 0 has been lost, i.e. $0 \in [\Omega]$. Thence we have to differentiate the following cases :

15. I) A certain text of Ω is known to be 0. Then all other texts are dependent on 0.

II) No text of Ω is known to be 0.

a.- $0 \in \Omega$: the elimination by single faults leads to 0.

b.- $0 \notin \Omega$, i.e. $0 \in [\Omega]$: the elimination by single faults does not leave any text. Then we have to search for the position of the archetype in (Ω, —).

The archetype itself contains faults (or it has to be treated as the original); some of them may have been corrected, whereas readings of the original cannot have been corrected. Thus in the case of 15, II) b.-, the elimination by single faults is misleading for two reasons :

I) In fact, there occur corrections of readings not only of the other texts but also of the archetype. Their number will be rather small; it may be supposed, however, that this number is not equally distributed among all texts of Ω, but that only a few copyists were able and inclined to correct the text and that particularly these same scribes would copy the text carefully, producing rather few faults. Thus the corrections will concentrate in some of the lists of variants and in these lists it will be more probable that a correction in the copy seems to be an error of the independent text.

II) Even if it were granted that copyists only add faults, the observed false reading may be an error of a text from which one of two main families of (Ω, —) derives. Then this family does not derive from any text of the other family, but both families derive from a text (archetype) which had not appeared in (Ω, —).

To avoid misleading conclusions, it is necessary here to keep close to the empirical facts of tradition, and it may be allowed that I refer to the material which I have gathered in my edition of the Anonymus Bellermannianus (7). There I have given complete lists of variants of 20 mss.

(one of them (ρ) not preserved), excluding only the lists of the two mss. dependent on (ρ) and generally excluding slight orthographical variants which would not produce other forms and words or allow to conjecture them. The following table contains the number of those readings (of words) in the named mss. which differ from those of their antecedents. The former must be equal in value or better or worse than the latter.

ms.	degradations	readings of equal value	ameliorations	ameliorations in percentage of worse + better
A	20		0	
Mon215	31		4	
NeapIIIC1	21		1	
Mut173	39		1	
BerolPhill1555	4		0	
Par2532	24	9	10	29
Par2458	10		1	
VatRoss977	7	0	7	50
Par2460	35		4	
VatBarb265	5	1	3	38
Ambr700	6		1	
Vat221	11		0	
LaurAcqu64	4 (8)	0	3	43
NeapIIIC5	24		1	
Vat1364	51		2	
B	47		3	
VatUrb77	31		1	
C	58		2	
D	4	5 (9)	3	43 or 50
(ρ)	37 (10)	3	10	

It is obvious that the ameliorations are concentrated in some mss. and amount there to a considerable percentage, up to 50 % in VatRoss977 and perhaps also in D. According to this table, therefore, only if the worse readings of a ms. X comprise more than 50 % of the not equal readings of the complete list of variants will it be sufficiently certain that the compared ms. Y does not depend on X. This does not determine whether or not X itself depends on Y, because both of them might depend on an unknown intermediate text. For additional assurance the maximal percentage of ameliorations should not be limited by 50 % of the not equal variants.
I propose to assume that in a copy the number of the ameliorations may be as much as double the number of its degradations, a proportion which can easily be tested.

Let $v_c(x_i)$, the value of the reading x_i in comparison to c_i, be the probability that x_i rather than c_i is correct ($v_c(x_i) + v_x(c_i) = 1$) and

${}^{x}_{y} := |\{ i : v_y(x_i) > v_x(y_i) \}|$ the number of readings in X which are

more likely to be correct than the corresponding readings of Y (= number of readings in Y which are worse than the corresponding readings of X, i.e. the inequality can be read from both sides). Then the formulae for the *test of independence* are

16. $${}^{x}_{y} > 2\,{}^{y}_{x} \Rightarrow X \nleftarrow Y$$ and

17. $$1/2\,{}^{y}_{x} \leqslant {}^{x}_{y} \leqslant 2\,{}^{y}_{x} \Rightarrow X \nleftarrow Y \vee X \leftarrow Y$$

Under the condition of 16, all chains branching from Y and not including X depend on X or a text between X and Y and need not be further considered.

Under the condition of 17, two cases have to be differentiated :

I) $1/2 \frac{Y}{X} \leqslant \frac{X}{Y} \leqslant 2 \frac{Y}{X}$ can be reduced to $\frac{X}{Y} > 2 \frac{Y}{X}$ by excluding all variants of little significance and taking account, e.g., only of those omissions which can hardly be restored, and which are not corrections of dittographies or eliminations of words held to have been originally marginal glosses.

II) The reduction to $\frac{X}{Y} > 2 \frac{Y}{X}$ is impossible.

Then we have to move to the contiguous direct enchainment with the longer list of variants until the condition of 16 is fullfilled, if necessary by the reduction of I).

18. The last relation X — Y to be examined can have one of two types of forms :

I) One of its texts has no further contiguous direct enchainment, i.e. $W — X — Y \not{—} Z \wedge W \nleftrightarrow X$ where X and Y may be interchanged, or

II) both X and Y have contiguous direct enchainments, i.e. $W — X — Y — Z \wedge W \nleftrightarrow X \wedge Y \nleftrightarrow Z$.

In both cases the result of the examination of the list of variants between X and Y can be a.- $X \nleftarrow Y$, b.- $X \nrightarrow Y$, or c.- $X \nleftarrow Y \vee X \leftarrow Y$.

I) a.- $W \nrightarrow X \wedge X \nleftarrow Y \Rightarrow$
$W \leftarrow [R] \rightarrow X \rightarrow Y \vee W \leftarrow X \rightarrow Y \vee$
$W \leftarrow X \leftarrow [R] \rightarrow Y$

b.- $W \nrightarrow X \wedge X \nrightarrow Y \Rightarrow$
$W \leftarrow X \leftarrow [R] \rightarrow Y \vee W \leftarrow X \leftarrow Y$

c.- $W \nrightarrow X \wedge (X \nleftarrow Y \vee X \leftarrow Y) \Rightarrow$
$W \leftarrow [R] \rightarrow X \rightarrow Y \vee W \leftarrow X \rightarrow Y \vee$
$W \leftarrow X \leftarrow [R] \rightarrow Y \vee W \leftarrow X \leftarrow Y$

II) a.- $(W \nrightarrow X \wedge Y \nleftarrow Z) \wedge X \nleftarrow Y \Rightarrow$
$W \leftarrow [R] \rightarrow X \rightarrow Y \rightarrow Z \vee$
$W \leftarrow X \rightarrow Y \rightarrow Z \vee$
$W \leftarrow X \leftarrow [R] \rightarrow Y \rightarrow Z$

b.- $(W \nrightarrow X \wedge Y \nleftarrow Z) \wedge X \nrightarrow Y \Rightarrow$
$W \leftarrow X \leftarrow [R] \rightarrow Y \rightarrow Z \vee$
$W \leftarrow X \leftarrow Y \rightarrow Z \vee$
$W \leftarrow X \leftarrow Y \leftarrow [R] \rightarrow Z$

c.- $(W \nrightarrow X \wedge Y \nleftarrow Z) \wedge (X \nleftarrow Y \vee X \leftarrow Y) \Rightarrow$
$W \leftarrow [R] \rightarrow X \rightarrow Y \rightarrow Z \vee$
$W \leftarrow X \rightarrow Y \rightarrow Z \vee$

$$W \leftarrow X \leftarrow [R] \rightarrow Y \rightarrow Z \quad \vee$$
$$W \leftarrow X \leftarrow Y \rightarrow Z \qquad \vee$$
$$W \leftarrow X \leftarrow Y \leftarrow [R] \rightarrow Z$$

The ambiguities arising from the texts which do not appear in (Ω,—), i.e. the question whether such a text should be assumed or not, sometimes can be solved. It may be assumed that there is a certain minimum and maximum number of degradations in a copied text (4 and 58 in the Anonymus-mss.). Thus the worse readings of a text will be no sign of dependence if the proportion of worse to better is more than ca. 1 : 15 (in the Anonymus-mss.) or, for more assurance, 1 : 20 (11), i.e.

19. $\genfrac{}{}{0pt}{}{y}{x} > 20 \genfrac{}{}{0pt}{}{x}{y} \quad \Rightarrow \quad X \leftarrow Y$

With $1/20 \genfrac{}{}{0pt}{}{x}{y} \leqslant \genfrac{}{}{0pt}{}{y}{x} \leqslant 20 \genfrac{}{}{0pt}{}{x}{y}$ it cannot be decided whether a lost text has to be assumed or not, lest there be special signs proving a direct dependence or excluding a correction. But such a detailed examination of not only ca. two thirds but of the whole list of variants is necessary only in the case of the last direct enchainment (before it was decided by one of the next steps), and here it is important too, because here it may be decided whether the archetype is lost or not.

From the evaluation of a reading x_i in X in comparison to the corresponding reading c_i in the collation-text C may be infered the evaluation of x_i in comparison to a reading y_i in Y which itself has already been valued in comparison to c_i. If both x_i and y_i differ from c_i and both are better or both worse than c_i, the relative value of x_i and y_i cannot be infered :

20.

	$v_c(x_i) < v_x(c_i)$	$v_c(x_i) = v_x(c_i)$	$v_c(x_i) > v_x(c_i)$
$v_c(y_i) < v_y(c_i)$	undecidable	$v_x(y_i) < v_y(x_i)$	$v_x(y_i) < v_y(x_i)$
$v_c(y_i) = v_y(c_i)$	$v_x(y_i) > v_y(x_i)$	$v_x(y_i) = v_y(x_i)$	$v_x(y_i) < v_y(x_i)$
$v_c(y_i) > v_y(c_i)$	$v_x(y_i) > v_y(x_i)$	$v_x(y_i) > v_y(x_i)$	undecidable

If x_i or y_i have not been valued, the inference is impossible. The positive difference of value may include those cases where there is no real amelioration but where there is evidence enough that such an amelioration has been intended.

Université de Göttingen

Dietmar NAJOCK

NOTES

(1) Paris, 1968.

(2) Prof. MAU and Mr. AHNERT have worked out programs for this purpose and intend to describe them in this periodical.

(3) If we could exclude secondary changes of variants (e.g. corrections), we might define the direct enchainment of X and Y by

$X — Y :\Leftrightarrow (X_C \subset Y_C \vee X_C \supset Y_C) \wedge \neg \exists Z_C : X_C \supsetneqq Z_C \supsetneqq Y_C$
with $X_C := \{ x_i \epsilon X : x_i \neq c_i, c_i \epsilon C \}$ being the set of variants of X in comparison to C.

(4) In Ω may appear direct enchainments which are not direct in Ω, i.e.
$X — [Z] — Y$ subenchainment of $(\Omega, —)$, $X,Y \epsilon \Omega$
$\Rightarrow X — Y$ subenchainment of $(\Omega, —)$.

(5) If a word contains elements of different model-words, these elements have to be regarded as terms.

(6) This is realized by the programs of Prof. MAU and Mr. AHNERT.

(7) Diss. Göttingen 1970.

(8) Not counting intended omissions.

Extrait de la Revue (R.E.L.O.)
VIII, 1 à 4, 1972. C.I.P.L. - Université de Liège -

(9) One perhaps better.

(10) But 11 hard to prove.

(11) It may happen, however, that the proportion of the minimum and maximum number of degradations is less than 1 : 20, but this case will be rather exceptional.