

AFFILIATION PROGRAMS (PART II)

USING THE PROGRAM "AFFILI"

An affiliation program working with variants of a certain text in differing copies as data and generating a diagrammatic representation of the genealogical connections of these copies has been worked out by Prof. J. MAU of the Classics Dept. of Göttingen University during the years 1969-1971. (1). This program based mainly on the work of Father FROGER (2) is called AFFILI. It was first written in FORTRAN IV for an IBM 7040 computer, later it was translated into FORTRAN V for UNIVAC 1108 (3).

The complete processing of establishing a stem is the following :

1. Collation of the texts
2. Preparing the collation-material for the computer
3. Generating and sorting of "constellations" and "constellation-types" by program
4. Transformation of the table of constellation-types into a table of "text-enchainment" by programs
5. Interpretation of the output
6. If there are various possible interpretations : exclusion of certain parts of the collation-material and repetition of steps 3 - 5
7. Orientation of the text-enchainment

Steps 3, 4 and 6 have already been treated by Prof. MAU (1), step 7 by Dr. D. NAJOCK (4) and in shorter form by Prof. MAU. Steps 1, 2 and 5 will be treated here and the processing of steps 3 and 4 will be shortly repeated in diagrams. The terminology in the following will be mainly the

same as in the quoted articles (1), (4).

1. *Collation of the texts*

- 1.1 The extant texts, both mss. or printed editions, are compared with a fixed collation-text C (see Prof. MAU : Expl. to tab. 2).
- 1.11 If C is the original or a text having been corrected by philological conjectures all differing readings are "mistakes" and the resulting text-enchainment of the output will show the historical genealogy of the texts; otherwise there are "variants" and the text-enchainment has to be orientated to find a historical stem (see step 7).
- 1.12 The collation-text C is understood as a 2-dimensional matrix, where each sign of the text (letter, figure, punctuation mark, blank) is associated with a pair of numbers : with the number of the line and the number of the sign in this line. The two numbers are written as one decimal fraction. It is possible to compose the first number out of a number of the page and a second one of the line (see Prof. MAU : Expl. to tab. 2).
- 1.2 The collation-text must be written in machine-readable form : punched cards, punched tape, magnetic tape etc.
Then it is reprinted for each collator, the numbers above mentioned automatically being associated with the single signs of the text.

Table 1

Example of a printout of the collation-text :

sign → 1 2 3 4 5
 line ↓
 313 SIBI IPSI CONVERTITUR . UNIVERSALIS QUOQUE AFFIRMATIVA
 314 DUPLICI CONCLUSIONE CONTINETUR . NAM QUISQUIS OSTENDIT
 315 A TERMINUM DE OMNI C TERMINO PRAEDICARI , ILLUD

- 1.21 It will be useful to print the numbers of the columns after every three lines of the text. The distances between the lines will be arranged convenient to the wishes of the scholar doing the collation.
- 1.22 Later on the wrong readings of this text can be replaced by those of better ones (see Prof. MAU), and thus the output can serve for the final edition.
- 1.3 Although the collation is often done by various persons, all texts have to be compared according to the same principles. Therefore we have tried to find a form as simple as possible for the description of text-variants.
 For any differing reading of a compared text at least the following information must be given (further information added as commentary is not read by the program AFFILI) :
- 1.31 a) the *place of the variant* : number of the line and first and last number of that sequence of signs in C which differs in the compared text
 b) the *different reading itself*
- 1.311 Simple variants :
- 1.3111 A single sign differs :
 Example : in line 313 a compared text reads "QUAQUE" for "QUOQUE". We write : 313.39 *A*
- 1.3112 More than one sign differs :

- Example : for the same "QUOQUE" another text reads "QUOD".
We write : 313.40-42 *D*
- 1.3113 Addition in comparison with C :
Example : the compared text has "ILLE" inserted between "CONVERTITUR" and the following full stop. We write the last identic sign (here a blank) followed by the added signs (notice the blanks before and after punctuation marks) :
313.22 *ILLE*
- 1.3114 Omission in comparison with C :
Example : in the compared text "QUOQUE" does not appear.
We write : 313.37-43 **
inscribing no entry between the asterisks.
- 1.312 Changes in the word-order :
as the "differing reading itself" (see 1.31) we write the place whereto the words or signs are transposed (using a "/" for the first "**").
Example : in line 313 "SIBI IPSI" is changed to "IPSI SIBI" in a compared text. We write : 313.01-05 /313.06-10*
The same would be : 313.06-10 /313.01-05*
but we need not denote this inversion.
- 1.313 Damaged readings :
The place is described as shown in 1.311. Every character in C the equivalent of which in the compared text cannot clearly be read is marked :
1) by the sign "-", if we can see that the reading of the compared text cannot have been that of C
2) by the sign "+", if it cannot be excluded that the reading of the compared text has been the same as in C.
- 1.3131 Example : in line 314 at the place of "NAM" remainings of a

- word of at least 6 illegible characters can be seen.
 We write : 314.34-36 *-----*
- 1.3132 Example : in line 315 in the place of "ILLUD" a word of equal length has been totally erased. We write : 315.43-47 *++++*
- 1.32 Concerning the *significance of a variant* we distinguish :
- 1.321 Variants of little significance, with a high probability of having been made independently from one another or having been corrected ex ingenio (e.g. orthographic variants, punctuation marks etc.). The significance of these variants is denoted by the number "1". They will in most cases be excluded when establishing the text-enchainment.
- 1.322 Variants of great significance; all those which can hardly have originated independently from one another or which can hardly have been corrected. Their significance is given a grading of "2".
- 1.323 If the collating scholar is not able to decide between "1" and "2", he will give no grade of significance. These variants will be considered in the first machine-run, but they may be excluded for a second one, if the generated enchainment will be non-informative.
- 1.33 *Additions to the basic text* are denoted. Such additions can be omitted to generate the text-enchainment, but they can also be used for a separate run of the program AFFILI.
- 1.331 The type of an addition is marked by a letter :
 "I" inscriptions, titles
 "G" glosses
 "S" scholia; others may be added.
- 1.332 The place of an addition is marked by the line-number (and, if necessary, by the column-number too) and :

“+” if the addition is above the line
“-” if the addition is under the line
“A,B,C,D” for the upper, right, lower, and left margin.

- 1.34 A *label of the text* compared with the collation-text C is given by one or two letters. Further hands in the same ms. are distinguished by an additional number from 1 to 9.
- 1.4 Texts compared according to other principles may in most cases be adapted to our form of description by a simple additional program (4).
- 1.5 All collation-data for a single as well as for several texts can be written on a printout of C or directly on punched cards, punched tapes, or magnetic tapes.

2. *Preparation of the collation-material*

- 2.1 Writing the collation-material in machine-readable form for use by the program AFFILI we have to observe the following *order of the single data* :
- 2.11 If necessary, a distinguishing mark for *additions to the basic text* : type and place of addition (see 1.331 and 1.332)
- 2.12 A decimal fraction denoting the *beginning of the variant-place*. The decimal fraction may have up to 10 figures, the number of figures must always be the same. Therefore in lower numbers add zeros on the left (see 1.31)
- 2.13 If the variant is longer than one character, the decimal-fraction denoting the *end of the variant* is given. The figures before the decimal point (and the point itself) can be left out, if they are

- identical with those of the beginning (see 1.31).
- 2.14 *Significance of the variant* (see 1.32)
- 2.15 *Label of the text* (see 1.34)
- 2.16 *The differing reading itself*, enclosed by asterisks (see 1.311). In the case of changed order of words we write for the first "*" a "/" (see 1.312).
- 2.17 Other *remarks* may be added then (e.g. paleographical or philological ones). They are not used by the program AFFILI, but they can be printed out for the scholar in a special output. The end of the remarks has to be indicated by "*". This asterisk has to be given, even if there are no additional remarks.
- 2.2 Between the single signs of a datum there may occur an arbitrary number of blanks, but of course not within a decimal-fraction, a text-label, or the differing reading itself (except those blanks which are part of the reading).
- 2.3 Collations standardized in other forms can be transformed by additional programs for AFFILI (see 1.4).
- 2.4 The formal correctness of collation-material standardized according to 2.1 will be tested by a special program that indicates those items which do not occur in the order described. The main-features of a testing program are already figured out by the author.

Table 2

Example of collation-data arranged according to 2.1 and 2.2 :

In addition to his explanations we have to take into account a further possibility not mentioned there :

- 3.2 If there occurs more than one variant at a single place of C and therefore more than one constellation is given for this place, these constellations are printed in a special output for a philological test.
- 3.21 If the variants differ only slightly, they are marked by a special sign, again given into the computer and then integrated into one "main-constellation".
Such a proceeding is necessary, since the significance of a variant has been stated in comparison with C only. The relative significance of different variants, however, has not yet been considered.
- 3.22 If the variants of one place differ from one another with a grading of "2" (see 1.32), it is impossible to integrate them into one "main-constellation".
These constellations remain separated and they are not used by the program AFFILI in the first machine-run. But later on it is possible to use them for a more complete interpretation of the text-enchainment.

Table 3

Example of forming constellations :
collation-data of various texts for one place of C are integrated into one constellation

1. collation-data for one place of C				2. transformed for forming constellations		3. constellations	
p	s	l	v	ABCDEF (b)	v	ABCDEF (c)	v
15.32-35	2	A	*ARUM*	011111	ARUM		
15.32-35	2	D	*ARUM*	111011	ARUM		
15.32-35	2	F	*ARUM*	111110	ARUM	011010	ARUM
15.32-35	2	B	*UM*	101111	UM		
15.32-35	2	C	*UM*	110111	UM	100111	UM
15.32-35	2	E	*ORUM*	111101	ORUM	111101	ORUM

Abbreviations : p = place of the variant
s = significance
l = label of the text
v = the variant itself
b = label given as a bit-string
c = constellation

Table 4

Example of transforming constellations into constellation-types :

1. constellations		2. constellation-types	
ABERSTU	variant	ABERSTU	frequency
0111111	A		
0111111	IT	0111111	2

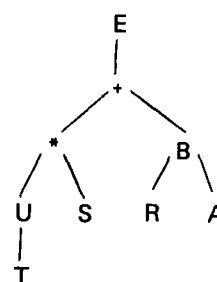
1110111	CEM.	1110111	1
1111011	NERA		
1111011	T	1111011	2
1111101		1111101	1
1111100	E	1111100	1
0010111			
0010111	TS		
0010111	N	0010111	3
1111000	NGE	1111000	1
0010000	/16-21	0010000	1
0000000	I		
0000000	STU	0000000	2

4. The “*table of constellation-types*” is generated by sorting the constellation-types in the program AFFILI. The same program transforms the “*table of constellation-types*” into the “*table of text-enchainment*”.

Table 5

Example of the output of the table of text-enchainment :

1. table of constellation- types	2. output of the table of text-enchainment	3. représentation of the stem in its usual form
ABERTU	texts	
111101	T	T U * + E
111100	TU	S * + E
111011	S	R B + E
111000	STU	A B + E
111011	R	
011111	A	
001111	AB	
001000	ABRSTU	
000000	ABERSTU	



5. *Interpretation of the output*

- 5.1 In the following we assume that the collation-text has been the original or an edition near to the original. Therefore we speak of "antecedents" in the genealogical chain.
If the collation-text was a text chosen at random, the output represents no genealogical stem but merely a text-enchainment (see 1.11). In that case the words "terminal, antecedent etc." are used cum grano salis.
- 5.11 The texts on the left of the computer-output of the table of text-enchainment (T,S,R,A) are "terminals", i.e. there exist no copies of these texts.
- 5.12 A text righthand of another one is his antecedent.
- 5.13 A text on the outmost right of a line is the oldest antecedent of all texts in this line.

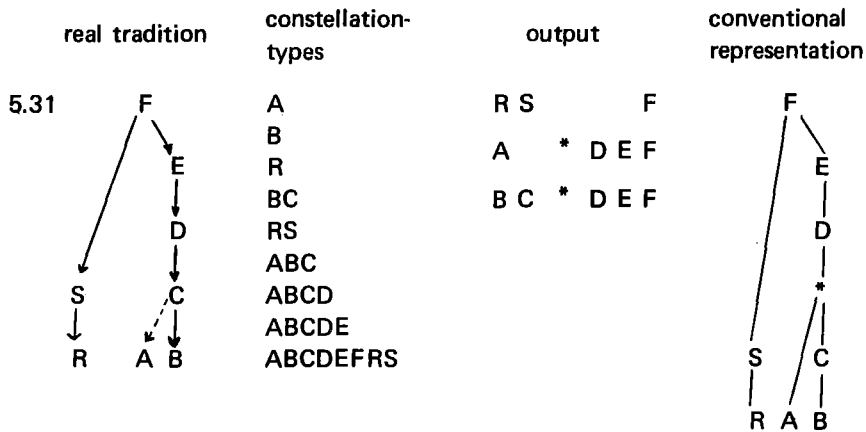
- 5.14 Lost texts (represented by "*" or "+") can be assumed only, if at least two extant texts have variants in common and if there is no further text which has only these variants.
Therefore the number of lost texts within the extant part of the tradition cannot be assumed greater than the number of extant texts minus 1. (In some cases of contamination or correction this may not be true.)
- 5.15 If a text in different lines of the output of the text-enchainment shows different antecedents, it is either contaminated or corrected.
- 5.2 In some cases of contamination and correction (see 5.3 and 5.4) there appear lost texts in the output. By merely logical methods it cannot be decided, whether they have really existed or whether their appearance is only due to the form of representation of the table of text-enchainment.
- 5.3 *Corrections* : If a scribe of a certain text has corrected a variant of an antecedent, the representation in the output differs according to whether the variant has been brought about by the direct antecedent or whether it has been done by a prior one.
- 5.31 If the variant that has not been taken over by our text occurs in the direct antecedent for the first time, both antecedent and correcting text appear depending on a lost text, which is inserted into the stem over the antecedent.
- 5.32 If some of the corrected variants have been brought about and some are taken over by the direct antecedent, a lost text appears in the output over any text that has brought about at least one of these variants and the very texts - except the oldest one - seem to be contaminated.
- 5.33 If all variants brought about by the direct antecedent have been

taken over, but some variants of a prior antecedent have been corrected, the representation of the stem will seem less obvious : Instead of the text that has brought about the variants not taken over by the correcting text a lost text appears in the output; and the connection between the direct antecedent of the correcting text and those descentents which have taken over all mistakes appears twice.

This kind of representation is due to a subroutine in the program AFFILI, and if interpretation will be too troublesome, it may easily be altered.

Table 6

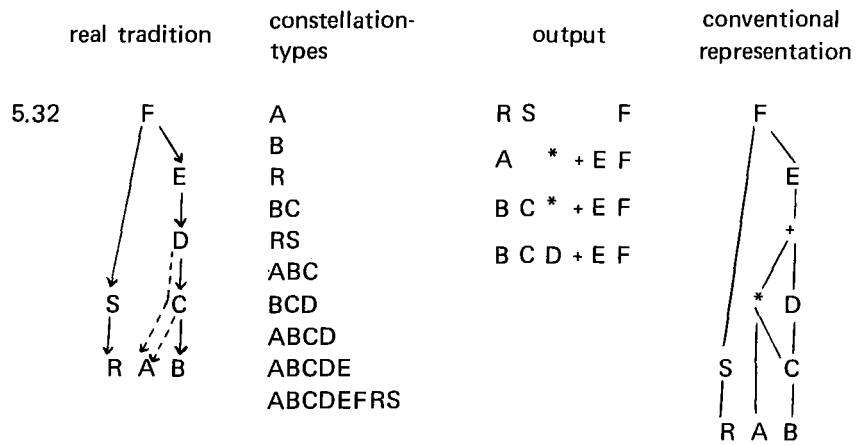
Examples of interpretation of the output in the case of correction :



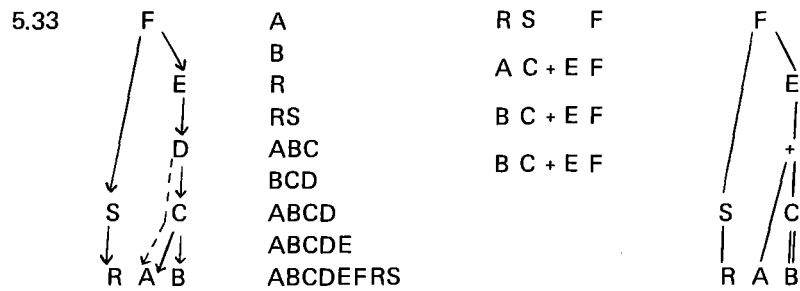
EXPLANATION : A has corrected some mistakes brought about by C

Continuation of table 6

Examples of interpretation of the output in the case of correction :



EXPLANATION : A has corrected some mistakes brought about by C and D



EXPLANATION : A has corrected some of the mistakes brought about by D. D does not appear in the output. In its place appears a lost text and a twofold connection between the direct antecedent of the correcting A and the descendents (here B) that have taken over all mistakes.

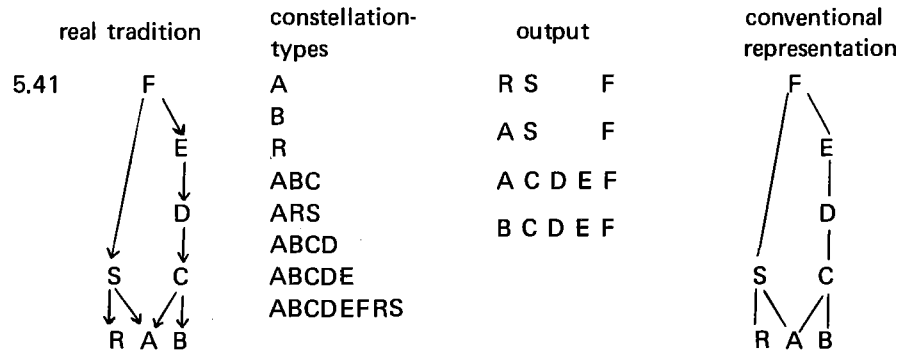
In the column of "real tradition" we have \rightarrow , if A has taken over all variants, and \dashrightarrow , if A has taken over only some but not all variants brought about by a certain text.

- 5.4 The interpretation of *contaminations* entails certain difficulties. Three fundamental types of contamination are to be differentiated :
1. The text has taken over all the variants of all its antecedents. This case is in reality very rare.
 2. The text has taken over all the variants of one main-antecedent and "additional variants" from other antecedents but never all their variants.
 3. The text has taken over variants from all its antecedents but never all their variants. We call the variants taken over by the contaminated text "optional variants".
- 5.41 If a text has taken over all the variants of all its antecedents, it appears in the output of the text-enchainment on the left of all these antecedents. There will be no special difficulty in interpreting this case.
- 5.42 If a text has additional variants outside a main-tradition, it is represented in the output once in the stream of the main-tradition and once as shown in the case of corrections (5.3), because in both cases the scribe has taken over some but not all variants of a tradition.
- 5.43 If there is no main-tradition and the scribe has chosen between optional variants, lost common antecedents of the contaminated text itself and of the sources of the variants not being taken over by the contaminated text appear in all branches of the stem on

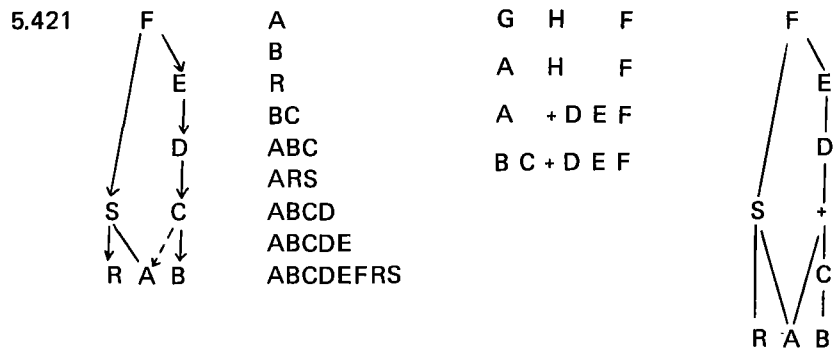
which the contaminated text depends. They are inserted into the stem according to 5.3.

Table 7

Examples of interpretation of the output in the case of contamination :



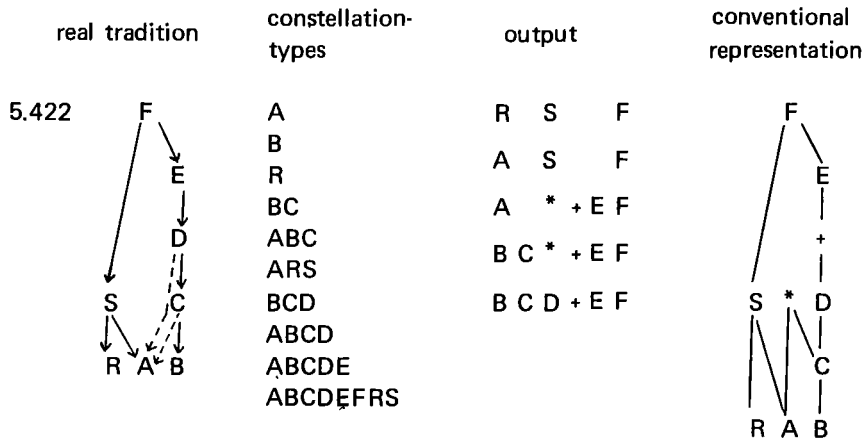
EXPLANATION : A has taken over all the variants of both S and C



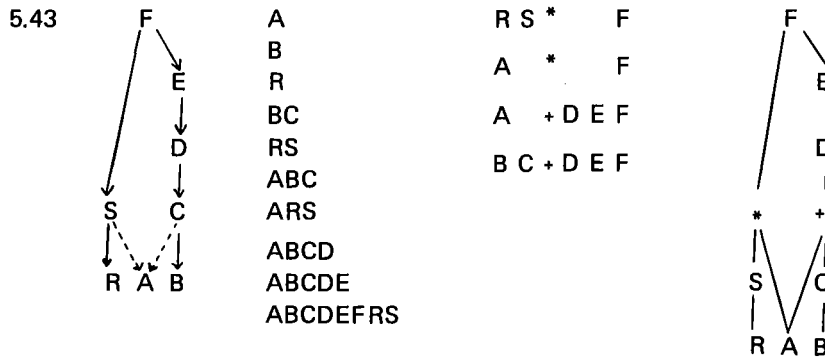
EXPLANATION : A has taken over all the variants of S and some of those brought about by C (but all those C had taken over from D)

Continuation of table 7

Examples of interpretation of the output in the case of contamination :



EXPLANATION : A has taken over all the variants of S and some of those brought about by C and some of those brought about by D (but all those D had taken over from E)



EXPLANATION : A has taken over some of the variants brought about by S and some of those brought about by C (but all those C had taken over from D)

ACKNOWLEDGMENTS

The author wishes to express his gratitude to Prof. Jürgen MAU who introduced mathematical methods and use of computers to the philologists of the University of Göttingen and who suggested and encouraged this work. He is also obliged to Dr. Dietmar NAJOCK who kindly offered to translate most of this article and suggested a lot of useful corrections.

D-34 GÖTTINGEN
Stegemühlenweg 31

Heinz-Jörg AHNERT

NOTES

- (1) MAU Jürgen, *Affiliation programs*, in *Revue de l'Organisation internationale pour l'Etude des Langues anciennes par Ordinateur*, 1972, n° 3.
- (2) FROGER D.J., *La critique des textes et son automatisation*, Paris, 1968.
- (3) FORTRAN V provides logical operations with the single bits of a computer-word and therefore is more convenient for handling big quantities of short informations. For smaller computers, however, AFFILI can be retranslated into FORTRAN IV. Most of the subroutines have been written for an IBM 1130 computer too, some in 1130-FORTRAN, some in 1130-ASSEMBLER. Apart from these programs of Prof. MAU the author tried to work out an affiliation program in 1130-FORTRAN using arithmetical operations instead of logical ones. This experiment proved that an affiliation program processed in such a small computer - at our disposal was an IBM 1130 with 8 K 16-bit words and single disk drive - takes too much time to be realized.
- (4) NAJOCK Dietmar, *Orientation of text-stems*, in *Revue de l'Organisation internationale pour l'Etude des Langues anciennes par Ordinateur*, 1972, n° 2, pp. 39-56.
- (5) We have to thank Prof. D.C. ANDRESEN and Dr. RITTER, editors of Dionysius the Areopagite, who placed collation-material of mss. of the letters of Dionysius at our disposal. This material collated according to principles being much more complicated than ours was transformed into our form of description by a program of the author.