

TECHNICAL INFORMATION EXCHANGE

IBM®

September 14, 1965

SORTING ALGORITHMS AND OPERATING
PROCEDURES FOR CARD SORTING ON THE
IBM/360 MODEL 20 WITH MFCM 2560

This paper describes, compares, and evaluates four different card sorting algorithms, giving extensive examples, tabulations and timing charts.

This paper also presents certain systems aspects and operating procedures which make system/360 card sorting more attractive than conventional card sorting.

Mr. W. N. Holmes
IBM Australia Pty. Limited
Box 88, P. O. St. Kilda
Victoria, Australia

For IBM Internal Use Only

CONTENTS

	<u>Page No.</u>
Abstract	1
Introduction	2
Straight Merge	4
Retrospective Merge	5
Compressive Merge	7
Long Input Strings	10
Synchronisation	12
Variable Algorithms	14
Reduction in Sorting Time	16
Continuous Operation	18
Reversed Files	20
Comparison with Unit Record Sorting	22
 <u>Appendices</u>	
Index	24
Section A - Merge Examples	25
Section B - Logic Diagrams	39
Section C - Test Results	42
Section D - Graphical Summary	46
Section E - Test Program	49

TITLE: Sorting Algorithms & Operating Procedures
for Card Sorting on the IBM/360/20 with
MFCM 2560.

AUTHOR: W. N. Holmes

DATE: December 30, 1964

DIRECT ENQUIRIES TO: W. N. Holmes,
IBM Australia Pty. Limited,
P.O. Box 88,
St. Kilda, Victoria, Australia.

ABSTRACT: This paper describes, compares and
evaluates four different card sorting
algorithms, giving extensive examples,
tabulations and timing charts.

This paper also presents certain systems,
aspects and operating procedures which
make System /360 card sorting more
attractive than conventional card sorting.

INTRODUCTION

This paper was inspired by several intra-office discussions on the utility of the 360/20/2560 for card sorting and the efficiency of the supporting program.

The investigations carried out to support this paper were hastily executed and the results included in the appendices must not be regarded as anything but flimsy evidence, based as they are on small samples called in an informal manner from the middle of the Smiths in the Melbourne Telephone Directory.

The algorithms investigated were merely the first few which sprang to mind and many more remain even unmentioned. In particular, digit sorting similar to conventional card sorting can be performed with the MFCM, but this has been ignored.

The casual reader will be mainly interested in the practical conclusions drawn in the section entitled "Comparison with Unit Record Sorting".

The paper will be most easily followed if the appendices are separated from the body of the paper for easy reference while reading the narrative.

INTRODUCTION (continued)

Relevant IBM Publications are:-

C26-3601

IBM System/360 Model 20
Punched Card Utility Programs

F28-8001

General Information Manual
Sorting Methods for IBM DP Systems

STRAIGHT MERGE

The concept of merging to create a file with a single sequence is quite distinct from that of digit sorting as with an 082 card sorter. It is assumed that the reader is familiar with both concepts.

The straight merge simply feeds two halves of an input file (one half from each hopper) and merges, string for string, then flip/flop stacks output, string by string, into two stackers. The process is exemplified in Appendix A1, with a logic diagram in Appendix B1.

Basically, the gain G of the process will be 2.0, the gain being defined so ---

$$G = \frac{\text{No. of } \overset{\text{input}}{\text{output strings}}}{\text{No. of input strings}}$$

However, with small input strings, the gain may rise slightly above 2.0. This adventitious process is illustrated in Appendix A2.

Using a straight merge, 2N input strings may be merged into a single sequence with N passes of the file.

RETROSPECTIVE MERGE

To improve the gain available with straight merging, the circumstances of adventitious gain (as shown by Appendix A2) may be amplified. The value of the control field of the last card selected into EACH stacker is relevant. This information is available to the controlling program and it would be a pity not to use it.

A "retrospective" algorithm might be loosely stated as follows. Feed the lower or only "possible" card into the highest "possible" stacker. If neither card is "possible", feed the lower card into the highest stacker.

By "possible" card is meant a card that can be stacked without causing sequence break. Also, a "possible" stacker will not break sequence if the card is selected into it. The ranking of stackers is by control field of the last stacked card. The algorithm as stated is for ascending sequence sorting.

The retrospective merge is exemplified in Appendix A3, with a logic diagram in Appendix B2. The example of Appendix A3 may be compared with that of App. A1. The retrospective merge is compared with other merges in Appendix A4.

RETROSPECTIVE MERGE (continued)

The appendices referred to above show input strings of average length 2. The process whereby a gain larger than 2 is achieved is more clearly illustrated in Appendix A6 which illustrates a retrospective merge with input string length of about 5.

One characteristic of the algorithm for retrospective merging is that its scope is not limited to a mere two stackers. The merge illustrated in Appendix A5 shows the merge of Appendix A3 but using three stackers. The tabulated test results of Appendix C4 give an indication of the dependence of gain on the number of stackers used.

COMPRESSIVE MERGE

In seeking to improve the retrospective merge, its operation must be examined in detail. Consider the merge illustrated in Appendix A3. Notice that card 731 in the primary feed is held (waiting for the 500 and 583 cards to be added to sequence in stacker 2) despite its close proximity to the 709 card in stacker 1. This card would have been held up even if it had been a 710 card. The 731 card is, in fact, followed by cards 504 and 558 which might advantageously be merged with cards 500 and 583 in the secondary feed.

The retrospective merge is designed to feed as many cards as possible without causing a sequence break. Taking a more indirect approach, it might be expected that a merge designed to minimise the gap in current stacker sequences, regardless of sequence breaks, would achieve a higher gain, at least with short input strings.

Hence, an algorithm for "compressive" merging might be stated as follows. Calculate the "gap" for each feed in combination with each stacker. Select the feed/stacker combination which gives the smallest gap and feed that card into that stacker.

COMPRESSIVE MERGE (cont.)

The gap across a sequence break is calculated according to the modulus of the control field. For example, with a positive three digit numerical field, the gap between 999 in the stacker and 001 in the feed is 002. With alphabetic fields the gap might be calculated by depth of match of the two fields, or by binary subtraction, or some such subterfuge. With /360 internal coding, the use of binary subtraction would tend to emphasise the sequence break, but not as utterly as retrospective merging. In fact, some experimenting with such a binary subtraction with all fields might yield some interesting effects, giving, as it would, a blend of retrospective and compressive merging.

The compressive merge is exemplified in Appendix A7, with a logic diagram in Appendix B3. Note that the = exit from the gap comparison was not implemented in the program used to generate the test results of this paper. The example of Appendix A7 may be compared with those of Appendices A1 and A3. Also, the tabulated test results of Appendix C2 show that the expected higher gain is definitely attained with short input strings.

The compressive, like the retrospective merge, is not limited to use with two stackers. Appendix A8 (cf. A5) shows the compressive merge using three stackers, and Appendix C4 tabulates test results showing the dependence of gain on the number of stackers used.

COMPRESSIVE MERGE (continued)

The appendices referred to above show input strings of average length 2. The process whereby a gain larger than 2 is achieved is more clearly illustrated in Appendix A9 which shows a compressive merge with input string length of about 5. Comparison of Appendices A9 and A6 clearly displays the essential difference in operation between retrospective and compressive merging.

LONG INPUT STRINGS

The results described so far show that sophisticated merges give gain significantly higher than a straight merge for short input strings. However, with long input strings, the compressive and retrospective merges degenerate to operation very similar to the straight merge. This effect will be seen in Appendix A10, which shows a retrospective merge with input string length of about 8. The similarity to a straight merge is immediately evident. The compressive merge tends to give similar results although cards may be fed in a slightly different sequence or large gaps in input sequence may hold up one feed for a time, giving very poor gain locally.

The extra gain achieved by the retrospective and compressive merges comes from the switching of the output from one stacker to another. At the point of switching, the control field in the stacker currently being fed approaches that in another stacker, and the gap in sequence allows the high stacker to take over the output, thereby reducing the gap which would occur by straight merging. In effect, each stacker has a switching point which (when using two stackers) steadily backs down the sequence until it crosses a sequence break, thus causing one less than the straight number of strings in one stacker.

LONG INPUT STRINGS (continued)

Hence, what might be called local gain approximately equal to one half the average gap will be obtained at each stacker switch. This gain should actually be greater than half the average gap since stacker switching might be expected to occur preferentially at the larger gaps. In commercial sorting applications cards usually cluster, and this will allow greater gain than a random distribution. In the extreme, groups of identical cards will reduce the effective string length since they will be fed and stacked as though they were a single card.

The use of more than two stackers will increase the gain possible since each stacker will have a switching point so that each output string will tend to spread itself over all available stackers. Appendices A11 and A12 illustrate this effect.

SYNCHRONISATION

Consideration of Appendices A11 and A12 reveals weaknesses in the operation of both retrospective and compressive when the input comprises long strings. The effect is that a blocking condition may occur to prevent one feed from contributing to the output until the block is removed at a significantly later stage. The effect can be so pronounced that merging of the input strings does not actually take place for some time and the controlling program only works one feed.

Appendix A11 shows that a high card, in conjunction with one or more low stackers, is a blocking condition for the retrospective merge. Appendix A12 shows that a card fed such that a large gap in sequence separates it from every stacker constitutes a blocking condition for the compressive merge.

The blocking effect may be prevented by forcing the input to merge, string for string, using the knowledge of the previously stacked cards to control the stacking of subsequent cards but not the feeding of the cards. The method might be thought of as synchronisation of the primary and secondary feeds.

Benefit from synchronisation would not be expected unless the average input string length were greater than the number of stackers used. Notice that application of this technique removes the distinction between retrospective and compressive merging, emphasising that the relative success of compressive merging at low string length is due to its better control over the cards fed, not from more efficient stacking.

Synchronisation was applied to the retrospective merge section of the program which produced the test results of the appendices to this paper. The results of synchronisation may be seen in Appendices A4, A13, C2, C3, C4, D1 and D2. In particular, the comparison in Appendix A4 between the synchronised retrospective merge and the straight merge demonstrates their similarity, at least on the input side.

The synchronisation described above is not the only method of removing the blocking effect, but it has the virtue of simplicity.

VARIABLE ALGORITHMS

The algorithms and variations already described are of varying effectiveness depending on such factors as input string length, stackers used and range and type of control field. This variability suggests that increased overall effectiveness could be obtained by keeping a measure of certain factors such as input string length (in each feed, perhaps) and switch from one type of merge to another when conditions appear to be favourable. For example, a weighted moving average string length or control field gap size might be kept (and revised for each card read to enable a switch to be made part way through a long string) or a range (perhaps weighted moving average again) for each digit or character of a field and/or for each field as an entity.

Appendix D3 gives an estimate of the performance a variable algorithm might achieve, compared with the performance of simpler merges. It must be emphasised that this appendix is based on rather insubstantial test results and extensive guesswork.

VARIABLE ALGORITHMS (continued)

Notice that retrospective and compressive merges may be interchanged at any point. If synchronism of the feeds is to commence or cease, certain points of interchange may be optimum. The straight merge would not be directly interchangeable unless only two stackers were in use throughout. However, when input strings become very long so there is no point in avoiding the straight merge, the extra stackers may be dropped one by one at sequence breaks, until a straight merge may be commenced.

REDUCTION IN SORTING TIME

The algorithms described in the preceding text are suggested as definite means to significantly reduce sorting time with the 360/20/2560. However, these are not the only means.

If a card file is being sorted into sequence solely so that it can be tabulated, then the final merge pass will best be written as part of the tabulating program to save one pass of the file. In fact, any such tabulating program should be written for the general case of two input files (including the special case, as say, an end of file run-out condition) which are incidentally merged.

Often a card file must be sorted to a series of related sequences (ringing the changes) to enable a series of related reports to be produced. With unit record sorting, the order in which the sorts are performed can be designed to minimise the number of card passes through the sorter by maximising the pre-sequencing. This can also be done with the 360/20 sorting, but the entire control field must be specified at every stage to force preservation of the pre-sequencing. However, the proportionate saving will probably not be as great.

REDUCTION IN SORTING TIME (cont.)

One incidental aspect of MFCM sorting is that mishandling of the cards will not invalidate the sort and indeed, if it occurs during an early pass, will probably go unnoticed.

Another aspect of MFCM sorting which promotes efficient operation is that the front of the file is not a significant point. In other words, the operator merely takes the cards from the front of the stackers and puts them in the hoppers (perhaps via card racks if the file is large) until they all feed into only one stacker. The file is then in sequence. There is no pause between passes to adjust the device. The aspect of continuous operation is examined more closely in the next section.

CONTINUOUS OPERATION

It might be thought that, with a straight merge, the procedure of taking cards out of stacker 1 (2) and putting them into hopper 1 (2) would ensure trouble-free continuous operation. This is not necessarily the case. For the first pass of the sort the operator will arbitrarily split the input file into two halves, one for each hopper. Since it would be very difficult to split the file initially into an equal number of strings, the merge would eventually fall into imbalance, the cards in one stacker/hopper loop building up, in the other disappearing. A little hand-merging of strings with pencil and paper will uncover this phenomenon. Hence the operator should eventually exercise control over the transfer of strings from stacker to hopper.

Also, when a more sophisticated merge is used, the problem of feeding two hoppers from say five stackers arises. For small string lengths, selection of large slabs from each stacker in turn will only introduce a few extra strings. However, when relatively few strings are left (if not earlier) the operator must exercise control over the transfer of strings from stacker to hopper.

CONTINUOUS OPERATION (Continued)

With a straight merge (either in its own right or as the last phase of a complex merge) the merge should split the strings between pairs of stackers. Within each pair of stackers, the feed would be switched at the first sequence break after, say, 500 cards. This would enable the operator to balance the hoppers without splitting any strings.

The use of marker cards (physically distinctive cards punched all 9's) is also recommended. These cards should be inserted at intervals throughout the initial input deck. They would indicate to the operator both the precise location of some sequence breaks and, indirectly, the progress of the sort and would be removed from the back of the deck on completion of the sort. In effect, the cards in any stacker between two marker cards would constitute a macro-string and the operator could replenish the lower hopper with the front or only macro-string in one of the stackers.

REVERSED FILES

Some reference is made in the IBM reference manual C26-3601-1 to the problem of reversing the sequence of a file from say descending to ascending order. Under any type of merge a strict reversal will be a lengthy procedure. However, if the file is predominantly in a sequence the reverse of that required for a report, a better shorter procedure is possible.

Sort, if necessary, the file to the strict reverse of the sequence required. Then, for the report, feed the file from the back face up nine edge first into the read hopper. This will supply input cards to the reporting program in the correct sequence but the card record will be inverted with columns 1 to 80 presented as columns 80 to 1. It is a relatively simple loop to invert the card using two registers for addressing, one incrementing, the other decrementing. The program requires only two instructions, on the other hand, if the translate special feature is installed. The coding used can be as follows.

First, some areas should be set up:

CARD	DC	CL80	CARD INPUT AREA
DRAC	DC	CL80	REVERTED CARD AREA

REVERSED FILES (continued)

Then a reformatting mask must be defined:

FMAT	DC	XL8'4F4E4D4C4B4A4948'
	DC	XL8'4746454443424140'
	DC	XL8'3F3E3D3C3B3A3938'
	--	-----
	DC	XL8'0706050403020100'

With the input record in DRAC the following instructions can be executed:

MVC	CARD, FMAT	SET MASK
TR	CARD, DRAC	INVERT

The inverted card image is now in the area CARD.

If it is anticipated that inverted files may occur within an installation, reporting programs may be written to read files optionally either normally or reversed. To determine if the file is in fact predominantly reversed in sequence, the average string length of a file (samples if large) can be determined and reverse sequence holds if the average is significantly less than 2. Alternatively, if the weighted moving average input string length calculated during a sort falls significantly (by some statistical measure) below 2, the sort can be halted by the program to allow reversed sorting to be substituted.

COMPARISON WITH UNIT RECORD SORTING

Two questions should be considered.

- (i) Should a card sorter be included in an installation?

If enough time for all necessary sorting is available on the /360/20/2560 without rental increase or overtime at suitable times of the reporting cycle, then there is no necessity for a card sorter. Advantages of simpler operating and reduced space requirements are gained by not including a sorter.

Notice that the 360/20/2560 would not be justified only on the basis of sorting, since great advantage in operating time will rarely be gained (more often lost) and it would be an expensive sorter. Operating time for sorts is discussed under the next question. The conditions within an installation may, on the other hand, prevent a card sorter from being justified.

- (ii) On which machine should a particular file be sorted?

It will not always be practical to sort a card file on the card sorter, for example if the control field contains binary or packed decimal data (summary cards) or is variable in format, or if a complex sequence determination is required. This last assumes that special sorting programs can be written within an installation because such programs would be small even if complex.

A sort performed on the /360/20/2560 will be more reliable since the necessary sequence checking will prevent "mis-sorting".

The time taken for a sort (exclusive of operator time) on either machine is proportional to the size of the file, the number of file passes and the feed speed. With a card sorter, the number of passes depends on the size and complexity of the control field. With the MFCM, the number of passes depends on presequencing and file size (hence double dependence) but is not dependent on the nature of the control field. Thus, for any particular sort there is a cutover file size, above which the MFCM is slower, below which faster.

The cutover file size, using the fairly sophisticated sorting algorithms and procedures of this paper, is roughly 5000 cards for a nine digit numeric sort, assuming random initial distribution and an 082 sorter. The effective cutover file size is raised (favouring MFCM) by complex control fields, operator time, presequencing and most other incidental factors. Appendix D3 will be found useful in estimating other cutover points, remembering that certain operating procedures described in this paper can reduce the number of passes below the number shown in that appendix.

Appendix A1 - Straight Merge.

INDEX TO APPENDICES

SECTION A MERGE EXAMPLES

1, 2 Straight Merge

3, 5, 6 Retrospective Merge

4 Comparison of Merges

7-9 Compressive Merge

10 Retrospective Merge (long input strings)

11 13 Merges using 5 stackers

SECTION B LOGIC DIAGRAMS

1 Straight Merge

2 Retrospective Merge

3 Compressive Merge

SECTION C TEST RESULTS

1 Explanation of Abbreviations

2 Sort Type v. Gain

3 String Length v. Gain

4 Stackers used v. Gain

SECTION D GRAPHICAL SUMMARY

1 String Length v. Gain

2 Input v. Output String Length

3 Passes v. Sort Type

SECTION E TEST PROGRAM LISTING

PRMRY	SECNY	ST.1	ST.2	PRMRY	SECNY	ST.1	ST.2
731				978		755	
	709			628*		978	
504*	500*	709			664		021*
	583	731		922			628
558			500	412*	095*		664
869			504		922	095*	922
	795		558	074*		412	
	211*		583		030*	922	
164*			795		142		030*
440		154*	869	936			074
	426	211			672		142
	158*	426		377*	066*		672
336*		440			195	066*	936
	746		158*		696		
563			336	829		377	
050*			563		494*	696	
	227*		746			829	
403		050*		145*			145*
	359	227		413			413
	336*	359		066*			494
		403			755		755
156*			156*		254*		
723			336	238		066*	
	765		723	054*		238	
858			765		295	254	
	284*		858		186*	295	
043*				883			054*
111		043*			850		186
793		111			136*		850
		284		141*			883
		793			862	136*	
732*			140*			141	
	644		644		811	811	
	993		732		394*		
342*			993			325*	
	551*				309*	862	
186*		342*		855			325*
	922	551		253*			394
	551*	922		066*			855
014*			186*			253*	
	676		551		702	309	
	558*		676		061*	702	
297		014*		057*	285		061*
551		297					066
735		551			523		285
		558			631		523
		735			881		631
753		753			045*		881
147*		784			261	045*	
	436*	784				057	
683			147*	084		057	
	062*		436	118		084	
982			683	531		118	
355*			982		452	261	
	345	062*			798	452	
	559	345		252*		531	
223*		355			755*	798	
	032*	559					252*
	065		032*	456			456
	644		065	723			723
791			223	960			755
	837		644		329*		960
337*			791	214*		214*	
	148*		837	530		329	
	137*	148*			728	530	
		337		134*		728	
147*			137*		450*		
	536		147	568			134*
007*			536		231*		450
	445*			402*			568
657		007*			979	231*	
	021*	445		138*		402	
755		657			732*	979	
PRMRY	SECNY	ST.1	ST.2	PRMRY	SECNY	ST.1	ST.2

Appendix A2 - Adventitious Gain
with straight merge

PRMRY	SECNY	ST.1	ST.2
882		060*	
529*	226*	353	
		882	
835	619		226*
		529	
	754		619
725*	675*		754
			835
516*	860	675*	
		725	
	269*	860	
213*	568		269*
			516
			568
959	209*	209*	
	238	213	
	393	238	
	400	393	
166*	389*	400	
255		959	
581			166*
	470		255
	297*		389
357*			470
			581
806	832	297*	
528*		357	
		806	
		832	
	056*		056*
	194		194
876	538		528
			538
	747		747
630*	185*		876
		185*	
672	092*	630	
619*		672	
	794		092*
352*		619	
		794	
543	735*		
205*		352*	
		543	
		735	
156*	081*		081*
	256		205
			256
	259		259
	119*		
	644	119*	
		156	
263		156	
522		263	
941		522	
	418*		418
560*		644	
		941	
497*	898		418
			560
	634*		898
638		497*	
		634	
	282*	638	
713		713	
466*			
	761		282*
266*			466
			761
899	496*	266*	
		496	
	792	792	
201*	934	899	
		934	
239	905*		201*

*One sequence derived
where 2 were expected!*

Appendix A3 - Retrospective Merge

PRMRY	SECNY	ST.1	ST.2	PRMRY	SECNY	ST.1	ST.2
731				412*		922	
	709				664		021*
	500*	709		074*			412
	583		500		695*		664
	795		583	936		074*	
504*					922	074*	
	211*	731			030*	095	
	426	795		377*			922
	158*	211*		829			936
		426		145*		377	
558		504				829	
869		558			142		030*
164*			869		672		142
	746		158*	413			145
440			164	068*			413
336*			440		000*		672
	227*	746		238		066*	
	359	227*			195	066*	
563		336			696	195	
	336*	359		054*		238	
050*			563		494*		696
403			050*		799	494	
	765	403			254*		755
156*		403		883			054*
	284*	765			106*		254
723		284*					295
	140*		723	141*		883	
858			858	811		141*	
043*			043*		850	141*	
111			111	394*		188	
793			793	899			811
	644				136*	394	
	593	644			862		850
732*		593		253*			855
342*		732		066*			136*
	551*	593			325*	862	
186*		342*			309*		253
	922	551		897*		066*	
	551*	551			702	309	
	476	551			061*		702
	558*	476			084		057*
014*			186*		285		061
	784		558				084
	436*	784		118			118
297		436*		531			285
551		297			523		
	062*	436		252*	431	929	
735		551				531	
753			735		881	631	
147*			753		045*	881	
	345		062*	456	261	045*	
			147			252	
683			345		452	261	
	559				798		452
	032*	559		723			456
982		683		960			723
355*		032*			799*		798
223*		582			329*	755	
	065	032*	355				960
	644	065		214*			214*
791		223		530			329
					728		530
	837				450*		728
337*			644			134*	
147*			791		568	450	
		337				568	
007*			837		231*		231*
	148*		147*				402
657			148		979		402
	137*			138*			
		007*			732*	979	
	536	137			387*		732
	445*		536				
	021*	445		335		138*	
755				105*		335	
978			657		762	387	
628*			755		620*		762
922			978		475*		620
		628					

Appendix A7 - Compressive Merge

PRMRY	SECNY	ST.1	ST.2	PRMRY	SECNY	ST.1	ST.2
731				628*	978		
	709			021*	445		
	500*	709		664	021*		
504*	583	731		922	095*	628	
			500		922	664	
558			504	412*	095	922	
869			558			922	
	795		583		030*	030*	
	211*	795			142	142	
164*		869			672	412	
440		164*		074*	936	074	
	426	211		377*	066*	672	
	158*	426			195	936	
336*		440			696	066*	
	746		158*	829	494*		195
563			336				377
050*		563					696
	227*	746					829
		050*					
403			403				
156*							
723							
	359	156					
	336*	227					
		359					
			723				
858			858				
043*			043*				
111			111				
793			336				
	765	765					
	284*	793					
732*			732				
342*							
186*	140*	284*					
	644	342					
014*			140*				
	993	644					
	551*	993					
		014*					
297			297				
551			551				
735			551				
			735				
753			753				
147*							
683							
	551*	147					
	676	551					
	558*	676					
		683					
982			982				
355*			355*				
223*			558				
	784	784					
	436*	223*					
791		436					
	062*		791				
337*			062*				
147*			337				
	559	559					
	032*	032*					
	065	065					
	644	644					
007*			134*				
	837	644					
	148*	837					
	137*		007*				
657			137				
	536	536					
	445*	657					
755			755				
978							
PRMRY	SECNY	ST.1	ST.2	PRMRY	SECNY	ST.1	ST.2

Appendix A8 - Compressive Merge using 3 stackers

PRMRY	SECNY	ST.1	ST.2	ST.3	PRMRY	SECNY	ST.1	ST.2	ST.3
731					731				
	709					445*	536		
	500*	709				922	628		
504*	583	731				412*	922		
			500			074*		412	
558			504				021*	445	
869			558				021*		
	795		583				074*		
	211*	795					095*	664	
164*		869					095		
440		211					030*	922	
	426	211					030*	936	
	158*	426					030*	936	
336*		440						030*	
	746		158*						
563			336						
050*									
	227*	746							
		050*							
403			403						
156*									
723									
	359	156							
	336*	227							
		359							
			723						
858			858						
043*			043*						
111			111						
793			336						
	765	765							
	284*	793							
732*			732						
342*									
186*	140*	284*							
	644	342							
014*			140*						
	993	644							
	551*	993							
		014*							
297			297						
551			551						
735			551						
			735						
753			753						
147*									
683									
	551*	147							
	676	551							
	558*	676							
		683							
982			982						
355*			355*						
223*			558						
	784	784							
	436*	223*							
791		436							
	062*		791						
337*			062*						
147*			337						
	559	559							
	032*	032*							
	065	065							
	644	644							
007*			134*						
	837	644							
	148*	837							
	137*		007*						
657			137						
	536	536							
	445*	657							
755			755						
978									
PRMRY	SECNY	ST.1	ST.2	PRMRY	SECNY	ST.1	ST.2	ST.3	

Appendix A9 - Compressive Merge

with input string length 5

PRMRY	SECNY	ST.1	ST.2	PRMRY	SECNY	ST.1	ST.2
297	858			045*	254	891	054
436		297		252		045*	
551		436		261			252
559		551		295			254
683		559		755			261
982		683		811			295
	043*	858		134*			755
	111		043	850			811
	140		111	855			850
	732		140	136*			855
032*		982		450		134	
065		032*		253		136	
223		065		325		253	
337			223	702		325	
007*			337	568		450	
	922		732	979		568	
	186*		922	138*			979
			007*	057*		702	
137		137		061			057*
445		186		084			061
	558	445		118			084
628		558		285			118
	735	628		452			138
922		735		387			285
	753	735		620			335
	062*	753		456			387
	147		062	723			452
	345		147	798			456
074*		922		960		723	
095		074*		214*		798	
377		095		788		960	
	355		345	802			620
	644		355	852			788
829		377		154*			802
	791		644	348			852
	837		791			154*	
066*		829		329		154*	
	147*		837	424		214	
	148		147	727		329	
	536		148	867		424	
195		066*		042*		530	
238		195		228		727	
494		238		388			867
883		494		435			042*
	657		536	402			200
	755		657	610			228
	978		755	669			388
141*		883		820			402
168			141	191		820	
394			168	475		105*	
	021*		978	597		613	
	412		021*	825		475	
862		394		632		597	
	664		412	761		613	
	922		664	951		632	
066*		671		240*			761
309		862		187			825
	936		922	482			951
	030*		936				187
	142		030*				482
	145						
	413		142				
523		309					
	672		145				
531		523					
631		413					
881		672					
	696		531				
	755		631				
	054*		881				
PRMRY	SECNY	ST.1	ST.2	PRMRY	SECNY	ST.1	ST.2

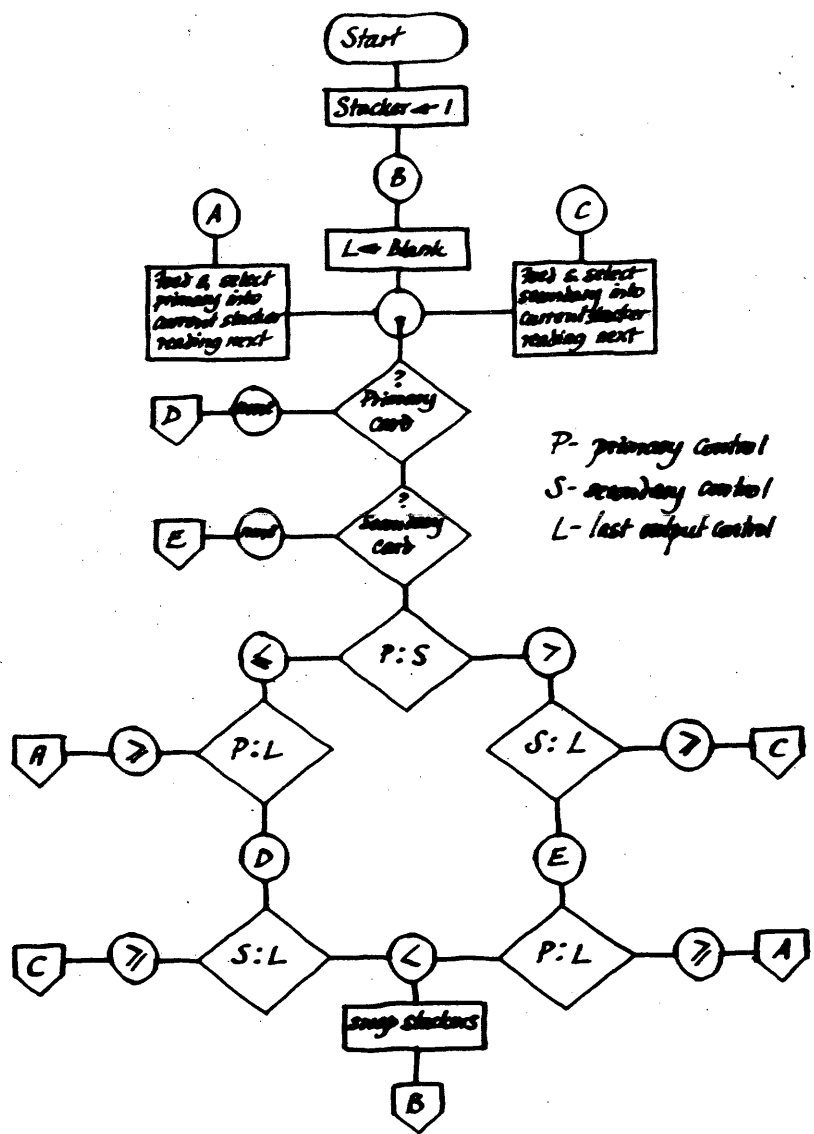
Appendix 10 - Retrospective Merge

with input string length 8

PRMRY	SECNY	ST.1	ST.2	PRMRY	SECNY	ST.1	ST.2
	165		153		806		698
	202		165	865		882	716
	206		202	CC3*		082*	806
	557		206	346		211	865
237			227	444		426	882
454			237	373		346	005*
551			454	444		440	082
568			551	483		644	211
	627		568	501			346
652			627	603			373
	641		651	860			426
	651		641	723			440
	662		652	858			444
	832		662	938			483
	821		749	121*			501
056*			821	043*			603
060			822	062			644
194			056*	111			723
	334		060	345			858
	213		084				938
	285		118				993
	335		138				043*
	387		285				062
	620		335				111
			387				121
			452				152
			456				282
			723				352
			798				536
			960				523
			214*				721
			788				551
			802				603
			852				791
			154*				523
			348				536
							551
							603
							683
							721
							763
							787
							791
							840
							866
							889
							972
							974*
							007*
							095
							142
							310
							479
							494
							377
							479
							494
							557
							648
							696
							709
							767
							811
							811
							855
							881
							928
							000*
							231
							C57
							311
							387
							419
							473
							613
							311
							315
							387
							419
							473
							651
							658
PRMRY	SECNY	ST.1	ST.2	PRMRY	SECNY	ST.1	ST.2

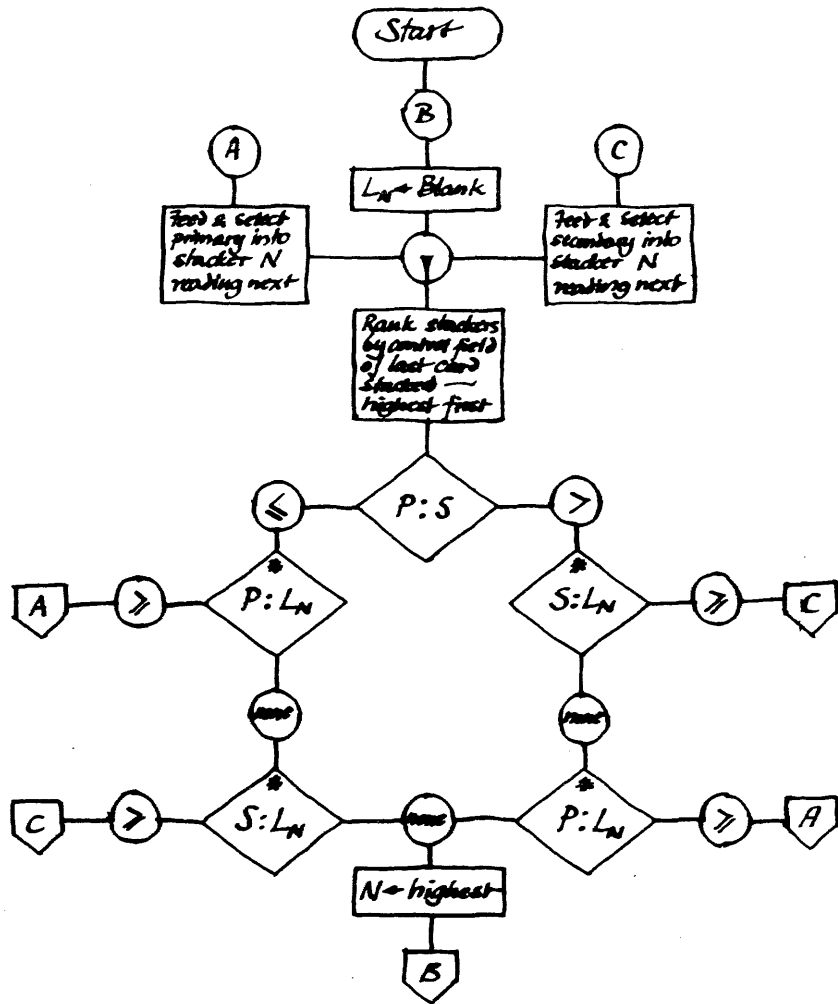
Appendix A13 - Synchronised Retrospective Merge using 5 stackers
 input string length 8

PRMRY	SECNY	ST.1	ST.2	ST.3	ST.4	ST.5
56C		543				
638		56C				
713	713	622				
85E		638				
	784				713	
	830				713	
	911				784	
859					830	
2C1*					898	
	C89*				899	
	128				911	
	194				089*	
	535				128	
239				194		
282				201		
287				239		
536				282		
	548		535	287		
616			536			
	579		548			
696			579			
782			616			
783	813	696				
144*		775				
		782				
	857	783				
	C85*	813				
	C57	857				
	223	C05*				
		C57				
16C						144
179						160
202						179
317						202
	270					223
	377					270
	388					277
581				317		
	402		388			
	617		402			
636			581			
708	644	617				
781	782	636				
980		644				
	545	708				
	193*	781				
051*		782				
085		945				
238		980				
	280	C51*				
345		C85				
	231	193				
	335	238				
	382					
429				331		
	562			335		
435				345		
591				382		
	682			429		
638				435		
828				562		
	684		591			
	E45		638			
115*			682			
	551		684			
	C86*		828			
	158		845			
			951			
		086				
PRMRY	SECNY	ST.1	ST.2	ST.3	ST.4	ST.5



Appendix B1 - Logic Diagram for Straight Merge/Split.

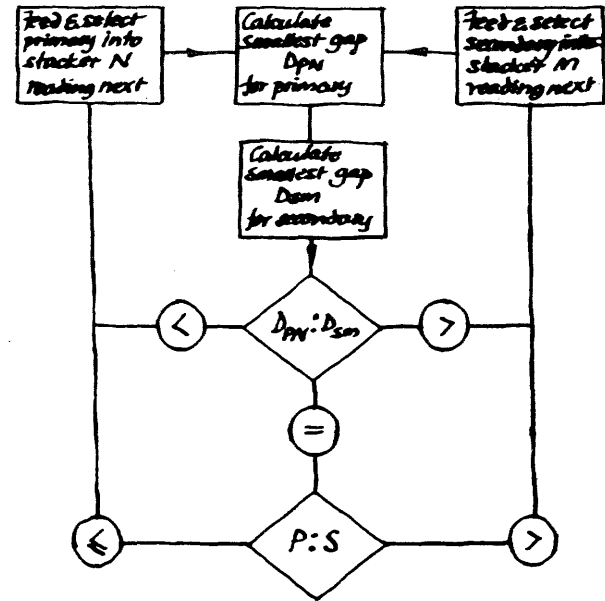
Appendix B2 - Logic Diagram for Retrospective Merge



Last Card logic is not included.

* - A search is made through established rank for the first stacker N which satisfies the '≥' exit condition.

Appendix B3 - Logic Diagram for Compressive Merge



Last card logic is not included.

43

N	Total number of cards passed
N _p	Number of cards through primary
N _s	Number of cards through secondary
S _p	Number of strings through primary
S _s	Number of strings through secondary
S ₁	Number of strings into stacker 1

S ₅	Number of strings into stacker 5
L ₁	Average length of input strings
L ₀	Average length of output strings
G	Gain = L ₀ /L ₁
T	Type of sort/merge
M	Straight merge/split
R	Retrospective merge
C	Compressive merge
R _s	Retrospective merge with synchronisation
K	Number of stackers used
X	Variable Merge

L ₁	T	G	L ₀	N	N _p	N _s	S _p	S _s	S ₁	S ₂	S ₃	S ₄	S ₅
2	R	3.9	8.1	802	387	415	184	206	20	20	20	20	19
	C	4.6	9.4	809	415	394	203	195	18	17	17	17	17
	R _s	4.0	8.1	827	416	411	204	204	21	21	20	20	20
5.3	R	2.8	14.8	754	364	390	67	75	11	10	10	10	10
	C	2.9	15.2	686	390	296	72	57	9	9	9	9	9
	R _s	3.0	15.7	768	390	378	72	73	10	10	10	10	9
8.4	R	2.3	19.3	772	382	390	45	47	8	8	8	8	8
	C	2.2	18.7	746	390	358	46	42	8	8	8	8	8
	R _s	2.4	20.3	773	390	383	46	46	8	8	8	7	7

Appendix C2 - Sort Type versus Gain
Stackers used = 5

43

45

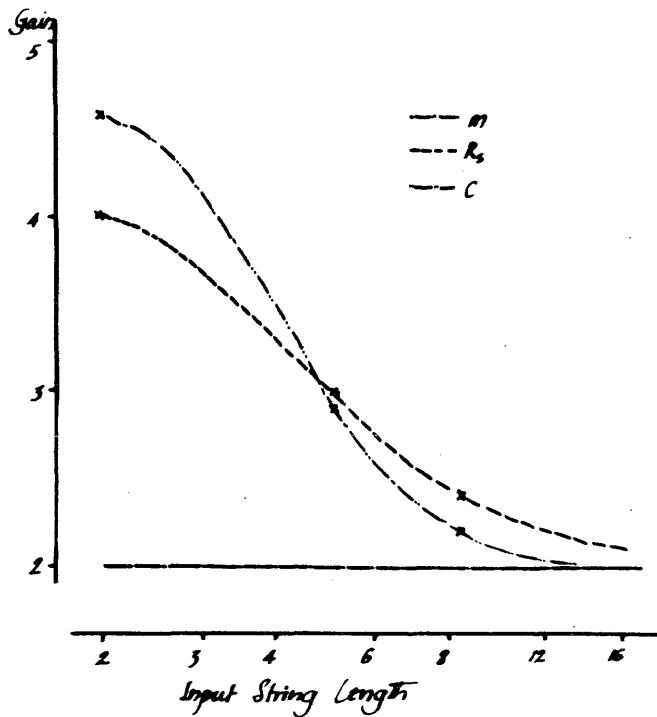
T	L ₁	G	L ₀	N	N _p	N _s	S _p	S _s	S ₁	S ₂	S ₃	S ₄	S ₅
R	2	3.9	8.1	802	387	415	184	206	20	20	20	20	19
	5.3	2.8	14.8	754	364	390	67	75	11	10	10	10	10
	8.4	2.3	19.3	772	382	390	45	47	8	8	8	8	8
C	2	4.6	9.4	809	415	394	203	195	18	17	17	17	17
	5.3	2.9	15.2	666	390	296	72	57	9	9	9	9	9
	8.4	2.2	18.7	746	390	356	46	42	8	8	8	8	8
R _s	2	4.0	8.1	827	416	411	204	204	21	21	20	20	20
	5.3	3.0	15.7	768	390	378	72	73	10	10	10	10	9
	8.4	2.4	20.3	773	390	383	46	46	8	8	8	7	7

Appendix C3 - Input String Length versus Gain
Stackers used = 5

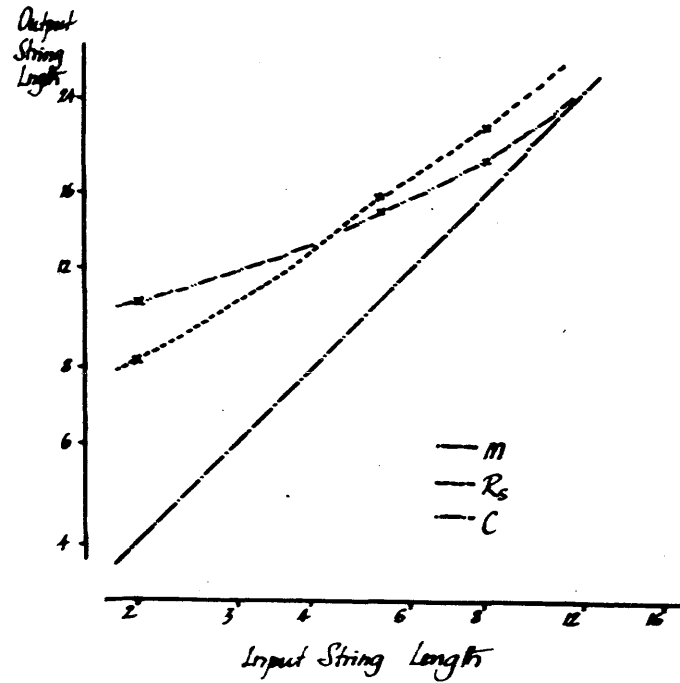
45

T	K	G	L ₀	N	N _p	N _s	S _p	S _s	S ₁	S ₂	S ₃	S ₄	S ₅
R	2	26	52	813	415	398	203	197	77	77			
	3	3.1	62	805	415	390	203	195	43	43	43		
	4	3.5	73	802	387	415	184	206	28	28	27	27	
	5	3.9	8.1	802	387	415	184	206	20	20	20	20	19
C	2	2.7	5.6	828	415	413	203	205	75	74			
	3	3.4	7.0	828	415	413	203	205	40	39	40		
	4	4.0	8.1	805	415	390	203	195	25	25	25	25	
	5	4.6	9.4	809	415	394	203	195	18	17	17	17	17
M	2	2.0	4.2	825	415	410	203	203	100	99			
R _s	2	2.5	5.1	827	416	411	204	204	81	81			
	5	4.0	8.1	827	416	411	204	204	21	21	20	20	20

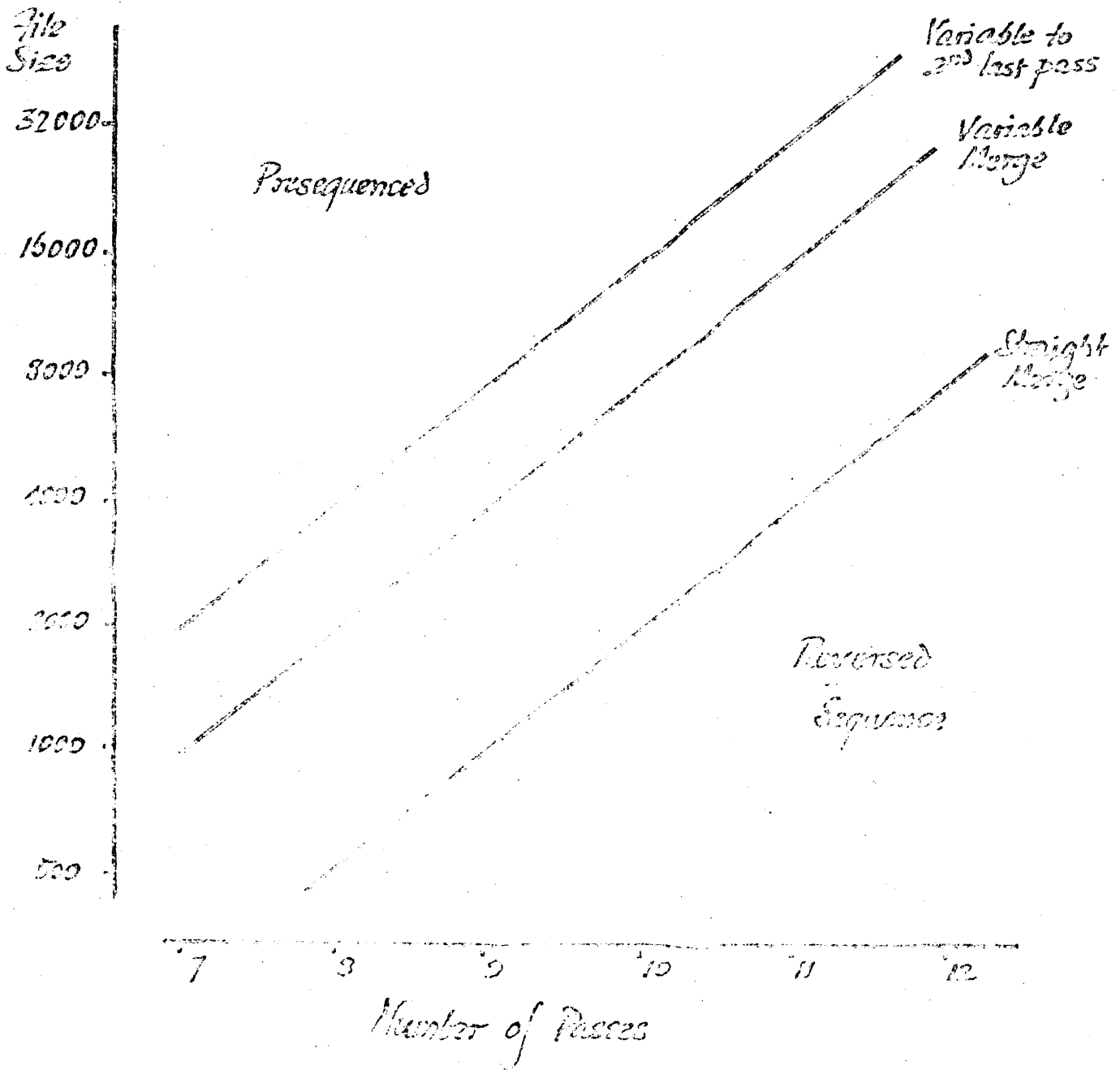
Appendix C4 - Stackers used versus Gain
Input String Length 2.



Appendix D1 - Input String Length versus Gain
Stackers used = 5



Appendix D2 - Input versus Output String Length
Stackers used = 5



Appendix D3 - File Sizes v. Passes

An appropriately drawn vertical line will give the correct point against card sorting 082.

APPENDIX E

The Sort Simulation Program

On the following five pages, the program used for deriving the results of this paper is listed with interspersed comments.

The program listing is given for interest & checking only, since it can hardly be presented as a model, having grown piecemeal over two or three days from a modest beginning with modest aims. It has no source program since it was punched directly into self-loading format.

The program used 3K of 1401, HLE compare, advanced programming, any 1403 and a 1402 with punch feed read. The loader, a standard one, uses modify address. Primary hopper input is to the read feed, secondary to the punch feed, of the 1402 card reader. A carriage tape with channel 1 punched is necessary, and 12 channel is interrogated for overflow. Input format is 26 three digit numbers per card in c.c. 1-78, the numbers being processed in turn from RIGHT to LEFT.

SORTING PROGRAM

CLEAR STORAGE & BOOTSTRAP CARDS

,008015,022026,030034,041,045,053L0721001026bbbbbb/0991,001/00111010?1bbbbbbC1
L096116,105106,110110B101/199,027#0990290027B001100bP026bbbbbbbbbcbbbbcbC2
,008015,022029,036040L071381,3543581001/080,001V00600611M467409,004424MbROOTLD01
L070412,366367,374381,386390,3973981001M468M003S474005#V358005KM007470bbLOADER02
L069442,406420,424431,439001,0010011001M46900000A47340904434690P374bbbLOADER03
L071474,447455,462466,468469,471474B350S469R424470M374420R406070bbb011bLOADER04

CONTROL ROUTINE FOR STRAIGHT MERGE

50150mbbbbmbbbbmbbbbmbbF0mF0m/180mC5035060P571TmC5095030P551TmPW01mP519bbbbbbbbb
551550C5095060BX01TmRW51mP5190C5095060P587TmP5630C5095030PX01TmP543bbbbbbbbb

ROUTINE TO GET NEXT PRIMARY NUMBER

701580H7340M8785030P7385030M8758780M7370P000mbbb/0800,001010L070070P705bbbbbb
0PRIMARY INPUT AREA INITIALISATION
87801bbbbbbbbbcbbbbcbbbbcbbbbcbbbbcbbbbcbbbbcbbbbcbbbbcbbbbcbbbbcbbbbcbbbbcbbbbcb

ROUTINE TO GET NEXT SECONDARY NUMBER

901300P9340MS785060P9385060MS75S780M9370mbbbbcbbbbcbbbbcbbbbcbbbbcbbbbcbbbbcb
931290Pbbbbmbbb/0800,001040L078S780P9350bbbbbbbbbcbbbbcbbbbcbbbbcbbbbcbbbbcb
0SECONDARY INPUT AREA INITIALISATION
57801bbbbbbbbbcbbbbcbbbbcbbbbcbbbbcbbbbcbbbbcbbbbcbbbbcbbbbcbbbbcbbbbcbbbbcb

CONTROL ROUTINE FOR RETROSPECTIVE MERGE

Y#148#F@#FJ#/18#B#J#1#C5#35#6#BY57#B#M#1#BY41#B#W#1#BY#9#B#N#1#B#P#1#####
Y4933#B#W51#B#Y#9#B#N#1#B#Y69#B#Y49#B#M#1#B#P#1#B#Y33#####

ROUTINE TO RANK THE STACKERS BY LAST FED NUMBER

J#148#H#K11#P#T13#Z13#H#Z15#T15#H#Z21#T21#B#J71#T25#H#Z27#T27#B#J71#T31#####
J4927#H#Z33#T33#B#J71#T37#H#Z39#T39#B#K#1#####
K#149#H#J#89#J#J#B#####Z/3#H#J#94#J#6#B#K88#Z#J#3#C#Z#J#8#Z/8#B#K6#J#T#H#J#94#!6#####
K5#J#5#B#K23#####M#Z#J#8#K59#M#M#Z/8#Z#J#8#M#M#K59#Z/8#M#B#K43#H#J#89#J#6#B#K#J#8#####
#STORAGE AREA FOR RANKED STACKERS
Z1331#####

LOOK FOR PRIMARY SLOT

M#154#H#J#99#H#J#89#J#J#B#J#J#Z/3#C#Z/85#J#B#M43#T#M#Z/5#W4#B#M#J#J#4#H#J#89#J#6#B#M12#####

LOOK FOR SECONDARY SLOT

N#154#H#J#99#H#J#89#J#J#B#J#J#Z/3#C#Z/85#J#B#M43#T#M#Z/5#W4#B#M#J#J#4#H#J#89#J#6#B#M12#####

FORCE STACKER WHEN NO SLOT

P#115#M#P14#Z18#B#K#1#####

SORTING PROGRAM - PAGE 4

CONTROL ROUTINE FOR COMPRESSIVE MERGE

Q0153MF@MFJm/180mBJ01mM503R03mBR04mM506R03mBR04mCQ88Q82mBQ62TmMC85W40bbbbbbbbbb
Q5435mBW01mBQ09mMQ79W40mRW51mBQ09mbbbmbbbmbbbmbbbmbbbmbbbmbbbmbbbmbbbmbbbmbbbmbbbmbbb
Q8912mBR04mRJ01mBQ24bb
00612mHQ23Q89mB350bb

ROUTINE CALCULATING GAP BETWEEN NUMBER & STACKER

R0156mbbbmHR67mMQ82Q88mMH089000mBR68Z/3+mCZ/8R03mBR93TmMZ/5Q79mSZ/8R03mbbbbbbbbb
R5750mMR03Q82mBbbbmH089000mS?06Z18mS?06Z18mBR43mH089000mBR23m500mbbbbbbbbbbbbbbbbb

GET NEXT PRIMARY NUMBER - MODIFIED TO SYNCHRONISE THE RETROSPECTIVE MERGE

70153mH753mM503759mM878503mR760503bmM875878mM756mC759503mR01TmRbbbbbbbbbbbbbbbbbb
75446mbbbmbbbm/0800,001m1mL078878mR712mPY41mBY57Tbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbbb
0134mH026784mR023Y24TmH026789mMbbBY24mR750mbbbbbbbbbbbbmbbbmbbbmbbbmbbbmbbbmbbb

GET NEXT 2NDRY. NUMBER - MODIFIED TO SYNCHRONISE THE RETROSPECTIVE MERGE

90153mH953mM506759mMS78506mR954506bmMS75S78mM756mC759506mR041TmRbbbbbbbbbbbbbbbbbb
95426m/0800,001m4RmL078S78mR912mPY60bb
04134mH066979mR063Y24TmH066789mMbbBY24mR950mbbbbbbbbbbbbmbbbmbbbmbbbmbbbmbbbmbbb

FORCE SYNCHRONISATION IN CONTROL ROUTINE

00616mHY32P01mHY04mR350bb

THE PRECEDING CARDS, CONCERNED WITH SYNCHRONISATION, ARE ONLY INCLUDED WHEN
WHEN THIS VERSION OF THE RETROSPECTIVE MERGE IS REQUIRED!

END CARD

006520B511B0047C0MT43T370047D0MT37T310B047E0MT31T250BY01G0BQ010bbbbbbbbbbbbbb

SENSE SWITCH SETTINGS

B ON - STRAIGHT MERGE & SPLIT
G ON - RETROSPECTIVE MERGE
ELSE - COMPRESSIVE MERGE

C ON - 5 STACKERS
D ON - 4 STACKERS
E ON - 3 STACKERS
ELSE - 2 STACKERS