

MEMBER NAME DOCGENSP

GENERAL SPECIFICATIONS

THE PURPOSE OF THE LONDON STAGE PROJECT INPUT SYSTEM IS TO READ IN DATA IN THE FORMAT DESCRIBED IN 'LOGICAL STRUCTURE OF INPUT DATA' AND, AFTER PERFORMING LADDER UPDATES ON 'AS', 'SEE', AND CAST GROUP LADDER ENTRIES, TO PRODUCE A DATA BASE CONSISTING OF SORT RECORDS WITH FIELDS RESERVED FOR THE FOLLOWING ITEMS ON EACH RECORD ..

- 1) DATE
- 2) SECTION TYPE
- 3) THEATRE
- 4) SYNTACTIC TITLE, IF ANY
- 5) TIME, IF ANY
- 6) SYNTACTIC ROLE, IF ANY
- 7) SYNTACTIC ACTOR, IF ANY

ALL POSSIBLE COMBINATIONS WILL BE PRODUCED IN THESE SORT RECORDS, THAT IS IF THERE IS A CAST GROUP WITH 2 ROLES AND 3 ACTORS THEN 6 SORT RECORDS WILL BE PRODUCED.

SCANNER INPUT DATA FORMAT

SCANNER INPUT IS ORIGINALLY IN THE FORM OF FIXED LENGTH BLOCKS 800 BYTES LONG, DIVIDED INTO 10 LINES OF 80 CHARACTERS EACH. THE LAST 5 CHARACTERS OF EACH LINE ARE SUPPOSED TO ALWAYS BE BLANK. THEY ARE TO BE IGNORED AND DELETED BY THE SCANNER INPUT EDITING PROGRAM, ICISCAN. DIFFERENT BLOCKING IS PERMISSIBLE AS LONG AS THE LRECL IS 80.

BATCHES OF INPUT DATA (ALL THE DATA INPUT TO ONE RUN OF A PROGRAM) MUST START WITH A PERFORMANCE AND MUST CONTAIN ONLY COMPLETE PERFORMANCES IN ORDER TO OBTAIN COMPLETE ACCURATE PROCESSING.

THE OUTPUT IS TO BE U FORMAT BLKS ACCEPTABLE AS INPUT TO A SYNTACTICAL PARSING PROGRAM CALLED STRCT (SHORT FOR 'STRUCTURE').

SINCE THE ICI SCANNER INTERPRETS THE DOUBLE QUOTE CHARACTER AS 2 SINGLE QUOTE CHARACTERS, ICISCAN WILL CONVERT ANY 2 CONSECUTIVE SINGLE QUOTE CHARACTERS IT FINDS INTO ONE DOUBLE QUOTE CHARACTSR. OTHER FUNCTIONS OF ICISCAN INCLUDE 'AT' SIGN PROCESSING (DEFINED ELSEWHERE), DELETING MULTIPLE BLANKS EXCEPT FOR 2 AFTER PERIODS, QUESTION MARKS, AND EXCLAMATION POINTS, CHECKING FOR ILLEGAL CHARACTERS, AND REBLOCKING INTO AN INTEGRAL NUMBER OF SECTIONS PER BLOCK, EXCEPT

IN ADDITION, DOUBLE BLANKS WILL BE RETAINED AFTER PERIODS, QUESTION MARKS, OR EXCLAMATION POINTS WHEN ANY OF THESE IS FOLLOWED IMMEDIATELY BY EITHER SINGLE OR DOUBLE QUOTES IN THE LOGICAL TEXT.

STARTING WITH THE FIRST, EVERY THIRD BLK OF SCAN INPUT MUST START WITH THE TYPIST'S INITIAL AND A PAGE NUMBER. THE ENTRY MUST BE THE VERY FIRST PRINTED CHARACTERS ON EACH TYPED PAGE, AND MUST BE IN THE FOLLOWING FORMAT ..

- 1) AN ASTERISK
- 2) A CAPITAL LETTER
- 3) 1 TO 5 DECIMAL DIGITS
- 4) A BLANK

TYPIST ENTRIES ON ON A LINE WILL BE PROCESSED BEFORE ANY 'AT' SIGN PROCESSING, SO TYPIST ENTRIES MUST BE CORRECT WITHOUT THE HELP OF 'AT' SIGNS. THIS SHOULD BE EASY SINCE ONLY THE FIRST FEW CHARACTERS OF THE PAGE ARE INVOLVED.

IN INPUT TO THE SCAN PROGRAM THE DELETIONS DEFINED BY 'AT' SIGNS WILL BE PERFORMED LINE BY LINE IN THE FOLLOWING ORDER ..

- 1) IF 3 CONTIGUOUS 'AT' SIGNS ARE FOUND ANYWHERE ON THE LINE THEN THE

MEMBER NAME DOCGENSP

ENTIRE LINE WILL BE DELETED FORTHWITH.

- 2) IF 2 CONTIGUOUS 'AT' SIGNS ARE ENCOUNTERED THEN THE 'AT' SIGNS AND ALL CHARACTERS BETWEEN THE 'AT' SIGNS AND THE FIRST BLANK TO THE LEFT OF THE 'AT' SIGNS WILL BE DELETED.
THE FIRST BLANK TO THE LEFT OF THE 'AT' SIGNS WILL NOT BE DELETED.
THIS WILL BE DONE LEFT TO RIGHT ACCROSS THE LINE REPEATEDLY
UNTIL NO DOUBLE 'AT' SIGNS ARE LEFT.
- 3) IF AN ISOLATED 'AT' SIGN IS THEN FOUND THEN THE 'AT' SIGN AND THE CHARACTER TO ITS LEFT WILL BE DELETED.
THIS WILL BE DONE LEFT TO RIGHT ACCROSS THE LINE REPEATEDLY
UNTIL NO 'AT' SIGNS ARE LEFT.

SUBTLETIES NOT DONE BY CHINA DATA ..

- 1) ABNORMAL (UNCERTAIN OR NONEXISTENT) DATES.
- 2) SOME BOX TAGGING, ESPECIALLY VERY ABNORMAL BOXES.
- 3) SUBTITLE TAGGING.
- 4) LADDER CHANGES EXCEPT FOR REPLACEMENT AND ADDITION AS IS MOST COMMONLY FOUND IN THE TEXT.
(THAT IS NO SIGNED LADDER CHANGES)
- 5) INDEX TAGGING
- 6) TIME NOTATION

0 ITEMGET SELECTION/EXCLUSION SPECIFICATIONS

THE ITEMGET PROGRAM WILL HAVE THE ABILITY TO SELECT INDIVIDUAL SORT RECORDS, GROUPS OF SORT RECORDS, SECTIONS OF SORT RECORDS, ENTIRE PERFORMANCES OF SORT RECORDS, AND TO EXCLUDE SORT RECORDS OUTSIDE OF CERTAIN DATE RANGES. ALL OF THE PRECEDING CAPABILITIES MAY BE COMBINED ALTHOUGH IT IS ENVISIONED THAT ONLY 1 OR 2 CAPABILITIES WILL BE USED IN ANY ONE RUN. SELECTION MAY BE BASED ON ANY COMBINATION OF THE ITEMS PRESENT IN A SORT RECORD. THESE ITEMS ARE .. SECTION TYPE LETTER, DATE, THEATRE, TITLE, TIME ENTRY, ROLE, AND ACTOR.

PROCESSING TO BE DONE FOR INDIRECT SORTS ..

- 1) MR., MRS., MISS., MLLE., SG., SGA., QUOTES, AND CAPITAL FOLLOWED BY A PERIOD WILL BE PLACED BEHIND THE ACTUAL NAME, INCLUDING COMBINATIONS OF THESE. SPACES WILL THEN BE TAKEN OUT AND ALL LETTERS CONVERTED TO UPPERCASE.
LEADING 'THE', 'A', AND 'AN' WILL BE TREATED SIMILARLY FOR TITLES.
SUFFIXES WILL NOT BE AFFECTED.

1 *** THE LOGICAL STRUCTURE OF INPUT DATA ***

THE LARGEST COMPONENT OF INPUT DATA IS THE S/360 LOGICAL RECORD, REFERRED TO HEREAFTER IN THIS REPORT AS A BLK (BLOCK). FOR SAVEDT INPUT EACH BLK SHOULD CONSIST OF ONE OR MORE COMPLETE SECTIONS, ALSO CALLED BOXES. EXAMPLES OF SECTIONS ARE .. PLAY SECTIONS, AFTERPIECE SECTIONS, DANCE SECTIONS, OPERA SECTIONS, SONG SECTIONS, ETC. THE SECTIONS ARE GROUPED INTO PERFORMANCES, WHICH CONSIST OF ONE PLAY SECTION (MAINPIECE PLAY) & ALL THE OTHER SECTIONS THAT FOLLOW THE PLAY SECTION AND PRECEDE THE NEXT PLAY SECTION. ANOTHER WAY TO THINK OF A PERFORMANCE IS 'ALL OF THE BOXES ASSOCIATED WITH ONE THEATRE ON ONE DATE'.

ALL OF THE ABOVE ALSO APPLIES TO SCANNER INPUT EXCEPT THAT THERE IS NO RESTRICTION ON SECTIONS BEING CONTAINED WITHIN ONE SCANNER INPUT BLK, ETC.

IT SHOULD BE NOTED THAT SOMETIMES A PLAY SECTION IS USED FOR SPECIAL PURPOSES OTHER THAN PROVIDING INFORMATION ABOUT A PLAY

MEMBER NAME DOCGENSP

GIVEN AT A CERTAIN THEATRE ON A CERTAIN DATE. THESE SPECIAL USES OF THE PLAY SECTION WILL BE DESCRIBED BELOW.

EACH SECTION STARTS WITH A THREE CHARACTER SEQUENCE THAT DEFINES BOTH THE START OF A SECTION AND THE END OF A PREVIOUS SECTION IF THERE WAS A PREVIOUS ONE IN THE SAME BLK. THE 1ST CHAR OF THIS TAG IS A BLANK. THE SECOND CHAR IS A UNIQUE CHARACTER THAT IS NEVER TO OCCUR FOLLOWING A BLANK AND IMMEDIATELY PRECEDING A SMALL LETTER ANYWHERE IN THE TEXT EXCEPT TO DEFINE THE START OF A SECTION. IT IS CALLED THE SECTION TAG CHARACTER. FOR MTST INPUT THIS CHARACTER IS THE PREFIX. FOR SCANNER INPUT AN ASTERISK IS USED. THE THIRD CHARACTER OF THE TAG IS A SMALL LETTER. ONLY CERTAIN LETTERS ARE ALLOWED. THESE ARE LISTED ELSEWHERE.

ALL SECTIONS EXCEPT COMMENT SECTIONS MAY CONTAIN 2 KINDS OF TEXT, STRUCTURED TEXT, AND EXTRANEOUS TEXT. A COMMENT SECTION CONSISTS ENTIRELY OF EXTRANEOUS TEXT. EXTRANEOUS TEXT CONTAINS ONLY ONE KIND OF ITEM OF SIGNIFICANCE TO PROGRAMMING, THE SIU TAGGED INDEX ENTRY. INDEX ENTRIES OCCUR ONLY IN EXTRANEOUS TEXT. THEY ARE DELIMITED ON THE LEFT EITHER BY A '\$' OR BY A '+', AND ON THE RIGHT BY AN EQUAL SIGN FOLLOWED BY 1 OR 2 TAG CHARACTERS. IF THERE ARE 2 TAG CHARACTERS THEN THE FIRST ONE IS A DECIMAL DIGIT. IF THE TAG IS A BLANK THEN THE TYPE OF INDEX ENTRY IS UNKNOWN. OTHERWISE THE TAG INDICATES SEMANTICALLY THE TYPE OF INDEX ENTRY, EG., ROLE, AUTHOR, STREET NAME, ETC. THE '\$' TYPE INDEX ENTRY INDICATES A PERSON OR PLACE .. THE '+' INDICATES AN ITALICISED INDEX ENTRY.

INDEX ENTRIES WITH A DOUBLE CHARACTER TAG MAY NOT EXCEED 125 CHARACTERS IN LENGTH. SIMILARLY INDEX ENTRIES WITH ONE-CHARACTER TAGS MAY NOT EXCEED 124 CHARACTERS IN LENGTH. BOTH THE ABOVE LIMITS INCLUDE THE INDEX DELIMITERS AND THE EQUAL SIGN.

FOR SAVEDT INDEX ENTRIES ALWAYS START WITH A BLANK-PREFIX-LARGE LETTER OR QUOTE. AN UNTAGGED OR UNKNOWN TYPE OF INDEX ENTRY CAN END WITH ANOTHER PREFIX-SMALL LETTER OR QUOTE. A TAGGED INDEX ENTRY ON THE MTST WILL BE INDICATED BY AN EQUALS SIGN AS THE RIGHT-HAND DELIMITER WITH THE TAG IMMEDIATELY FOLLOWING THE EQUALS SIGN. A BLANK IN THIS LOCATION IS ANOTHER WAY OF INDICATING AN UNKNOWN TYPE OF INDEX ENTRY, AS WITH THE SCANNER.

EXCEPT FOR COMMENT SECTIONS EXTRANEOUS TEXT IS DEFINED IN EITHER OF THE FOLLOWING 2 WAYS ..

- 1) ON THE LEFT A LEFT BRACKET, AND ON THE RIGHT BY ONE OF THE FOLLOWING ..
 - A) FIRST OCCURRENCE OF A PERCENT SIGN FOLLOWING THE LEFT BRACKET.
 - B) FIRST OCCURENCE OF A RIGHT BRACKET FOLLOWING THE LEFT BRACKET PROVIDED THAT ANOTHER LEFT BRACKET DOES NOT LIE BETWEEN THE ORIGINAL LEFT BRACKET AND THE RIGHT BRACKET.
- 2) ON THE LEFT WITH A LEFT PAREN, AND ON THE RIGHT WITH A RIGHT PAREN.

EXTRANEOUS TEXT DELIMITED IN ONE OF THESE 2 WAYS MAY CONTAIN DELIMITERS USED BY THE OTHER METHOD.. THAT IS, FOR EXAMPLE, EXTRANEOUS TEXT DELIMITED BY BRACKETS MAY CONTAIN PARENTHESES.

THE DELIMITERS ARE CONSIDERED PART OF THE EXTRANEOUS TEXT. EACH PAIR OF DELIMITERS MUST DEFINE EXTRANEOUS TEXT TO BE ENTIRELY WITHIN ONE SECTION. IF THERE IS NO MATCHING (RIGHT PAREN FOR LEFT PAREN, RIGHT BRACKET OR PERCENT SIGN FOR LEFT BRACKET) RIGHT DELIMITER FOLLOWING THE LEFT DELIMITER IN THE SAME SECTION THEN THE ENTIRE REMAINDER OF THE SECTION FOLLOWING THE LEFT DELIMITER WILL BE CONSIDERED EXTRANEOUS TEXT, HOWEVER AN ERROR MESSAGE WILL BE GENERATED.

PAGE ENTRIES ARE A WAY OF INFORMING THE COMPUTER OF WHAT PAGE OF THE LONDON STAGE ANY ITEM OF TEXT CAME FROM. PAGE ENTRIES ARE IDENTIFIED ANYWHERE IN THE TEXT BY A 'P' (BLANK-SMALL P) FOLLOWED IMMEDIATELY BY

MEMBER NAME DOCGENSP

AN UNSIGNED DECIMAL NUMBER WHICH IS ENDED BY A BLANK ON THE RIGHT ..
FOR EXAMPLE ' P132 '. NOTE THAT THIS PAGE ENTRY IS 6 CHARACTERS LONG
INCLUDING THE BLANKS, WHICH ARE IMPORTANT AND NECESSARY.

PAGE ENTRIES CAN OCCUR IN FRONT OF MOST ITEMS, HOWEVER THE FOLLOWING
RESTRICTIONS AND CONVENTIONS MUST BE FOLLOWED ..

- A) A PAGE ENTRY MAY NOT OCCUR INSIDE A DATE ENTRY, IT MUST EITHER
PRECEDE THE ENTIRE DATE OR ELSE PRECEDE THE THEATRE ENTRY WHICH
FOLLOWS THE DATE.
- B) SIMILARLY PAGE ENTRIES CANNOT OCCUR INSIDE A LADDER ENTRY.
- C) PAGE ENTRIES MAY NOT FOLLOW A 'BUT' - THAT IS 'BUT P312 ITEM'
IS NOT ALLOWED, EVEN THOUGH THE PAGE ENTRY DOES IMMEDIATELY
PRECEDE AN ITEM.
- D) THE PAGE ENTRY MUST IMMEDIATELY PRECEDE THE ITEM, SEPARATED FROM
THE ITEM ONLY BY ONE OR MORE BLANKS OR THE EQUIVALENT (CARRIAGE
RETURNS, LINEFEEDS, AND MTST RUBOUT CHARACTERS).
- E) NOTE THAT A SECTION DELIMITER SEQUENCE OR ANY OTHER TYPE OF
PUNCTUATION IS NOT AN ITEM.
- F) PAGE ENTRIES MAY NOT FOLLOW A TIME ENTRY. INSTEAD THEY
MUST EITHER PRECEDE THE TIME ENTRY OR PRECEDE THE ITEM
THAT FOLLOWS THE ITEM THAT FOLLOWS THE TIME ENTRY.

ROMAN PAGE ENTRIES, ANALOGOUS TO ORDINARY PAGE ENTRIES, ARE USED
IN EXTRANEOUS TEXT DERIVED FROM INTRODUCTIONS. IT IS SIMILAR
TO THE ORDINARY PAGE ENTRY IN THAT IT IS PRECEDED AND FOLLOWED BY
AT LEAST ONE BLANK. IT IS SYNTACTICALLY DEFINED BY A BLANK-EXCLAMATION
POINT COMBINATION IMMEDIATELY PRECEDING A SERIES OF ONE OR MORE
LEGITIMATE ROMAN NUMERAL CHARACTERS AND ENDING WITH A BLANK.
EXCEPT THAT IT IS ALLOWED ONLY IN INTRODUCTIONS, IT MAY OCCUR ANYWHERE
THAT AN ORDINARY PAGE ENTRY COULD OCCUR IN A COMMENT SECTION.

THERE ARE 4 BASIC KINDS OF SECTIONS ..

- 1) PLAY SECTIONS. THIS KIND OF SECTION IS REQUIRED TO HAVE
A THEATRE PRECEDING ITS TITLE. IT MAY ALSO HAVE ONE OR MORE
PARTS OF A DATE PRECEDING THE THEATRE.
- 2) SIMPLE TITLED SECTIONS DO NOT HAVE A DATE, THEATRE, OR TITLE TIME
ENTRY. ALL TITLED SECTIONS EXCEPT PLAY SECTIONS MAY
BE SIMPLE TITLED SECTIONS.
- 3) TITLE SECTIONS WITH TITLE TIME ENTRIES HAVE A TIME ENTRY PRECEDING
THE TITLE OF THE SECTION .. THE TIME ENTRY IS OF THE SAME FORM
AS FOR CAST GROUP TIME ENTRIES, THAT IS THEY MUST PRECEDE THE TITLE
IN A SECTION, AND ARE SEPARATED FROM THE TITLE BY THE COLON-BLANK
DEFINING THE END OF THE TIME ENTRY. WHEN A TITLE TIME
ENTRY IS ENCOUNTERED THE EFFECT IS ..
 - A) ALL GROUPS IN THE CAST LIST WITH AT LEAST ONE ROLE OR ACTOR
HAVE THE TIME ENTRY APPENDED TO THE GROUP.
 - B) IF THERE ARE NO ROLES OR ACTORS IN THE CAST LIST THEN A GROUP
CONTAINING ONLY THE TIME ENTRY WILL BE APPENDED TO THE END OF
THE CAST LIST.
- 4) UNTITLED SECTIONS DO NOT HAVE SECTION TITLES, DATE ENTRIES,
OR THEATRE ENTRIES. IF ANY OF
THESE SORTS OF ITEMS HAPPEN TO BE IN AN UNTITLED SECTION THEN
IT WILL BE TREATED AS SOMETHING IN A CAST GROUP.

THERE MAY BE MORE THAN ONE TITLED SECTION ONE ONE DATE AT ONE THEATRE
(IN ONE PERFORMANCE), BUT THERE MAY NOT BE MORE THAN ONE EACH OF ANY
UNTITLED KIND OF SECTION. THERE MUST BE EXACTLY ONE PLAY
SECTION, AND OF COURSE, IT COMES FIRST IN EACH PERFORMANCE.

THE FOLLOWING SECTION TYPES ARE SPECIFIED ..

P PLAY	PLAY	MUST HAVE THEATRE & TITLE, MAY HAVE DATE ENTRY .. NO TIME ENTRIES OF ANY TYPE ALLOWED AT ALL.
--------	------	---

MEMBER NAME	DOCGENSP	
A	AFTERPEICE	TITLED NO TIME ENTRIES OF ANY TYPE ALLOWED AT ALL.
C	COMMENT	EXTRANEous MAY CONTAIN SIU INDEX ENTRIES.
D	DANCE	UNTITLED
S	SONG	UNTITLED
M	MUSIC	UNTITLED
E	ENTERTAINMENT	UNTITLED
I	INSTRUMENTAL	TITLED
B	BALET	TITLED
T	TRICK	TITLED
U	MONOLOGUE WITH PARTS	TITLED
O	OPERA	TITLED

AN OPERA SECTION IS REALLY A SONG SECTION HAVING A SYNTACTICAL TITLE AND CAST LIST INSTEAD OF JUST A CAST LIST. IT WILL NOT BE DISTINGUISHED BY CHINA DATA FROM OTHER SONG SECTIONS.

ANY OF THE FOLLOWING IS TO BE CONSIDERED A GROUP OF STRUCTURED INPUT ..

PERFORMANCE HEADER - DATE, IF ANY, AND THEATRE (A PAGE ENTRY IS GENERALLY NOT CONSIDERED A PART OF THE GROUP).

TITLE OF A SECTION
 BALLET PART-BALLERINA/BALLERINO
 MUSICAL INSTRUMENT-MUSICIAN
 OPERA PART-OPERATOR
 TRICK PART-TRICKSTER
 ROLE-ACTOR GROUPS
 ENTERTAINMENT-ENTERTAINER GROUPS
 SONG-SINGER GROUPS
 DANCE-DANCER GROUPS
 MUSIC-MUSICIAN GROUPS
 'AS' TYPE LADDER REFERENCES
 'SEE' TYPE LADDER REFERENCES

FOR THE PURPOSES OF THE TYPIST ANY OF THE FOLLOWING THINGS IS AN ITEM ..

ROLE
 ACTOR
 TIME ENTRY
 SYNTACTIC TITLES
 SYNTACTIC SUBTITLES
 (PLAY, AFTERPEICE, OPERA, BALLET, TRICK, ETC.)
 OPERA
 BALLET
 INSTRUMENTAL MUSIC
 TRICK
 THEATRE
 EACH PART OF A DATE (BUT FOR SOME PURPOSES THE ENTIRE DATE IS CONSIDERED TO BE AN ITEM)
 LADDER REFERENCE
 INDEX ENTRY
 DANCE
 DANCER
 MUSICAL PEICE
 MUSICIAN
 SONG
 SINGER

MEMBER NAME DOCGENSP
ENTERTAINMENT
ENTERTAINER

THE UNTITLED TYPES OF SECTIONS (E,D,M,S) MAY EITHER CONTAIN A CAST LIST OR THEY MAY BE EMPTY, EXISTING ONLY TO SHOW THAT THERE WAS SOME OF THEIR PARTICULAR TYPE OF ACTIVITY ON THE GIVEN DATE AT THE GIVEN THEATRE.

A SYNTACTIC TITLE MUST ALWAYS BE PRESENT IN A TITLED SECTION.

IF THE TITLE IS NOT KNOWN THEN EITHER 'NONE' OR 'LADR' MUST BE PUT IN AS THE TITLE. IF SUCH A SECTION HAS A TITLE LADDER REFERENCE THEN 'LADR' IS USED, OTHERWISE 'NONE' IS USED. THESE ARE SPELLED WITH THE FIRST LETTER CAPITALIZED, THE REST SMALL. THEY ARE USED BOTH TO INDICATE THESE SPECIAL CASES AND TO SATISFY THE REQUIREMENT THAT EACH TITLED SECTION HAVE A SYNTACTIC TITLE.

THE SAME THING GOES FOR THEATRE ENTRIES IN PERFORMANCE SECTIONS, EXCEPT THAT 'NONE' IS ALWAYS USED, AND IT IS TYPED ENTIRELY IN SMALL LETTERS. SIMILARLY, IF ALL OR PART OF A DATE IS UNKNOWN THEN '0' MUST BE PUT IN FOR EACH UNKNOWN PART. THE VALUE OF EACH PART OF A DATE THAT IS WRITTEN AS ZERO REMAINS THE SAME AS IT WAS BEFORE, BUT THAT PART OF THE DATE IS MARKED IN THE OUTPUT DATA AS BEING ACTUALLY UNKNOWN. THE ENTIRE DATE THEN SERVES ONLY AS AN ESTIMATE OR GUESS FOR THE REAL DATE. SINCE NO PART OF A DATE THAT IS WRITTEN AS ZERO IS CHANGED, YOU MUST WRITE SUCCESSIVE UNCERTAIN DATES EXPLICITLY WITH 0'S IN THE UNCERTAIN PARTS OF THE DATE.

NO LADDER REFERENCE CAN REFER TO A PERFORMANCE THAT HAS AN UNCERTAIN DATE, THEATRE, OR TITLE.

LADDER REFERENCES ARE OF THE FORM ..

'AS/SEE XX MONTH (XXXX)'

WHERE 'XX' IS A NUMBER BETWEEN 1 AND 31 REPRESENTING THE DAY OF THE MONTH, 'MONTH' IS A 3 TO 5 CHARACTER STRING REPRESENTING THE MONTH OF THE YEAR, AND 'XXXX' IS A NUMBER REPRESENTING THE YEAR. THE PARENTHESSES AROUND THE YEAR INDICATE THAT THE YEAR PART IS OPTIONAL. IF THE YEAR IS NOT ENTERED THEN THE CURRENT YEAR FROM WHICH THE LADDER REFERENCE IS BEING MADE IS TAKEN TO BE THE YEAR REFERRED TO. BOTH 'AS' AND 'SEE' MAY START WITH EITHER LARGE OR SMALL LETTERS. A PERIOD MAY FOLLOW THE MONTH AND/OR YEAR PARTS, BUT NO OTHER PUNCTUATION, ESPECIALLY COMMAS, SHOULD APPEAR BETWEEN THE PARTS OF A LADDER REFERENCE.

NO SECTION MAY HAVE MORE THAN ONE TITLED TYPE LADDER REFERENCE IN IT. NO TITLE TYPE LADDER REFERENCE CAN BE PRECEDED BY ANY GROUPS IN THE SAME CAST LIST. OF COURSE IT MAY BE PRECEDED BY A TITLE, THEATRE, OR DATE ENTRY IN THE CASE OF A PERFORMANCE SECTION OR A TITLE IN THE CASE OF A TITLED SECTION.

THE END OF A LADDER ENTRY MAY BE DELIMITED EITHER BY A SEMICOLON, COMMA, MINUS SIGN, OR END OF SECTION, DEPENDING ON CIRCUMSTANCES.

'AS' TYPE LADDER REFERENCES WILL BE DONE AS DESCRIBED IN THE 5 MAY OCRB II 'TYPING INSTRUCTIONS FOR THE LONDON STAGE PROJECT'. THE LADDER CHANGE CAPABILITIES ARE TO BE ..

- 1) UNATTACHED ROLE SUBGROUP ADD
- 2) UNATTACHED ROLE SUBGROUP DELETE
- 3) UNATTACHED ACTOR SUBGROUP ADD
- 4) UNATTACHED ACTOR SUBGROUP DELETE
- 5) GROUP DELETE (PAIRED OR UNPAIRED)
- 6) ACTOR SUBGROUP & ROLE SUBGROUP REPLACE
- 7) ATTACHED ROLE ADD NO QUESTIONS ASKED

MEMBER NAME DOCGENSP

- 8) ATTACHED ROLE DELETE
- 9) ATTACHED ACTOR ADD NO QUESTIONS ASKED
- 10) ATTACHED ACTOR DELETE
- 11) PAIRED GROUP ADD / SAME FORMAT AS REPLACE

'SEE' TYPE TITLE LADDER ENTRIES

- 1) ALL FIRM GROUPS (THOSE THAT ARE NOT BROUGHT FORWARD BY THE LADDER REFERENCE) WILL REMAIN AS THEY ARE, ALL ITEMS AND GROUPS BEING CONSIDERED UNSIGNED.
- 2) A MATCH WILL BE ATTEMPTED ON EACH ROLE IN THE FIRM GROUPS. IF THE MATCHING ATTEMPT SUCCEEDS FOR ANY ROLE THEN THAT ROLE IS DELETED FROM THE QUESTIONABLE SET BEING BROUGHT FORWARD .. IF THE ROLE DELETED IS THE ONLY ROLE IN ITS GROUP THEN THE ENTIRE GROUP IS DELETED.

CAST GROUP LADDER ENTRIES

- 1) MUST IMMEDIATELY PRECEDE THE FIRST DASH IN THE GROUP IF THE GROUP HAS A DASH IN IT (BECAUSE EVERY ITEM FOLLOWING THE CAST GROUP LADDER REFERENCE IS A SYNTACTIC ACTOR, NO USE CONFUSING PEOPLE).
- 2) KEYING IS DONE ON ALL SYNTACTIC ROLES IN THE GROUP DOING THE REFERRING. IF SOME UNMATCHED ROLES ARE STILL LEFT OVER IN THE FIRST GROUP HAVING A MATCH IN THE SECTION REFERRED TO THEN NO ERROR IS GENERATED. IF ANY OF THE THE ROLES IN THE REFERRING GROUP DO NOT FIND A MATCH IN THE FIRST GROUP FOUND HAVING ANY MATCH IN THE REFERRED TO SECTION THEN AN 'INKY' ERROR MESSAGE IS ('INKY' MEANS 'INCOMPLETE KEYING IN GROUP IN WHICH SOME KEYING HAS ALREADY BEEN ACHEIVED') GENERATED. IF NO MATCHES AT ALL ARE FOUND THEN A 'NCGM' ERROR ('NCGM' MEANS THAT NO MATCH FOR ANY ROLE IN THE LADDER GROUP WAS FOUND IN THE REFERRED TO SECTION) MESSAGE IS GENERATED. THE SUCCESS OF THIS SCHEME DEPENDS ON NOT HAVING A SYNTACTIC ROLE MATCHING ANY OF THE ROLES IN THE REFERRING GROUP PRESENT PRECEDING THE ACTUAL GROUP ON WHICH THE KEYING IS DESIRED IN THE REFERRED TO SECTION.
RELATIVELY MINOR PROGRAMMING CAN SUBSTANTIALLY ELIMINATE THIS RESTRICTION.
- 3) ONLY 'AS' TYPE REFERENCES ARE ALLOWED.
- 4) TIME ENTRIES WILL BE BROUGHT FORWARD INTACT UNLESS THE CURRENT GROUP HAS ITS OWN .
- 5) ARE ALLOWED ONLY IN UNTITLED SECTIONS
- 6) AFTER SUCCESSFUL KEYING IS ACHEIVED, THEN ANY ACTOR IN THE REFERRED-TO GRP THAT IS MATCHED BY AN ACTOR PRECEDED BY A MINUS SIGN IN THE REFERRING GRP IS DELETED, ALONG WITH THE ACTOR PRECEDED BY THE MINUS SIGN IN THE REFERRING GRP, OF COURSE. IF THERE IS AN ACTOR PRECEDED BY A MINUS SIGN IN THE REFERRING GRP THAT HAS NO MATCH IN THE REFERRED-TO GROUP THEN A 'CGAD', 'CAST GROUP ACTOR DELETE' ERROR MESSAGE WILL BE GENERATED. (NOT YET IMPLEMENTED).
- 7) THE RESULT OF A SUCCESSFUL MATCH IS ALL OF THE SYNTACTIC ACTORS IN BOTH GROUPS, INCLUDING POSSIBLE REPETITIONS WITH NO QUESTIONS ASKED, EXCEPT FOR DELETIONS AS NOTED ABOVE. REPETITIONS MAY BE DELETED WITH RELATIVELY MINOR PROGRAMMING LATER.

1 PROGRAMMING SPECIFICATIONS

SPECIFICATIONS TO BE DEVELOPED ..

- 1) SYSTEM FOR TAGGING SUBTITLES TO BE USED IN LADDER KEYING.
- 2) SYSTEM FOR PATCHING, IF ANY.

MEMBER NAME DOCGENSP

NOTES ..

- 1) UNTITLED SECTIONS MAY HAVE 'TITLE' TYPE LADDER REFERENCES. NATURALLY, ONLY ONE UNTITLED SECTION OF ANY ONE TYPE IS ALLOWED IN ANY PERFORMANCE.
- 2) NO REPEATED TITLES, INCLUDING 'NONE', ARE ALLOWED IN ANY PERFORMANCE (FOR THE SAME TYPE OF SECTION).
- 3) IN LADDER UPDATES TIME ENTRIES WILL BE BROUGHT FORWARD INTACT UNLESS THE CURRENT GROUP ALREADY HAS A TIME ENTRY.
- 4) THE MAXIMUM BLKSIZE ALLOWED FOR ANY OF THE PROGRAMS IS GUARANTEED TO BE LESS THAN OR EQUAL TO 3625, HOWEVER THE DATA LENGTH OF BLOCKS IN THE LADDER PROGRAM IS ALLOWED TO BE AS SMALL AS 3593 BYTES INCLUDING CONTROL CHARACTERS. LARGER SECTIONS THAN THIS WILL REQUIRE SPECIAL EDITING AND SMALL TO MEDIUM SIZED CHANGES IN THE EDT, SAVEDT, STRUCTUR, LADDER, & ITEMGET PROGRAMS.
- 5) A SLASH WILL BE USED TO INDICATE THE END OF A LINE OF POETRY IN ALL CASES. A SLASH WILL BE USED TO INDICATE THE FIRST LINE OF A POEM. A SLASH HAS NO PROGRAMMING SIGNIFICANCE.
- 6) IN THE FUTURE ITALICS THAT DO NOT SIGNIFY TITLES WILL PROBABLY BE WIPE OUT.
- 7) TIME ENTRIES ARE NOT ALLOWED IN PLAY OR AFTERPIECE SECTIONS, BUT AN ERROR MESSAGE WILL NOT NECESSARILY BE GENERATED ON DETECTING A TIME ENTRY IN THESE SECTIONS.

FOR THE MTST SYSTEM THERE ARE 7 PROGRAM STEPS IN THE DATA ENTRY AND RETRIEVAL CYCLE ..

- 1) EDT
- 2) SAVEDT
- 3) STRUCTUR
- 4) LADDER
- 5) ITEMGET
- 6) SORT
- 7) FORMAT

A PATCH STAGE WOULD PROBABLY WORK BEST FOLLOWING THE LADDER STEP. PERHAPS PATCHES COULD BE IN A FORMAT LIKE THE LADDER CHANGES THEMSELVES.

*** DEFINITIONS ***

BALLET SECTION A TITLED SECTION TAGGED WITH THE SMALL LETTER 'B'. ITS DATA SHOULD COME FROM A DANCE BOX.

CAST GROUP IS A CHAR STRING IN THE CAST LIST PART OF A SECTION. IT IS DELIMITED ON THE RIGHT EITHER BY THE END OF THE SECTION OR BY A SEMICOLON. IT IS DELIMITED ON THE LEFT BY EITHER THE BEGINNING OF THE CAST LIST OR THE END OF THE PREVIOUS GROUP (BY A SEMICOLON).

CAST GROUP LADDER ENTRY A LADDER ENTRY IN AN UNTITLED SECTION THAT FOLLOWS A SYNTACTIC ROLE IN A CAST GROUP AND IMMEDIATELY PRECEDES THE FIRST DASH IN THE GROUP IF THE GROUP HAS A DASH.

CAST LIST ITEM TITLE LADDER ENTRY ('AS' OR 'SEE'), CAST GROUP LADDER ENTRY, CAST GROUP TIME ENTRY, SYNTACTIC ROLE, OR SYNTACTIC ACTOR.

CAST LIST ENTRY CAST GROUP.

CAST LIST TIME ENTRY AN ITEM IN A CAST GRP PRECEDING ALL ROLES AND ACTORS IN THE GRP. IT IS DELIMITED ON THE RIGHT BY BY A COLON-BLANK. IT CONSISTS OF THE COLON-BLANK AND USUALLY SOME OTHER CHARACTERS FROM THE SET, '0123456789ABDTIVX', WHERE A, B, D, AND T ARE SMALL LETTERS. ALSO CALLED A CAST GROUP TIME ENTRY.

EXTRANEous TEXT A) COMMENT SECTIONS

MEMBER NAME DOCGENSP

B) TEXT IN OTHER SECTIONS THAT IS DELIMITED AS EXTRANEOUS TEXT.

MONOLOGUE WITH PARTS A TITLED SECTION TAGGED WITH THE SMALL LETTER 'U'.

OPERA SECTION A TITLED SECTION TAGGED BY THE SMALL LETTER 'O'. ITS DATA SHOULD COME FROM A SONG BOX IN THE CALENDER.

PERFORMANCE ALL OF THE SECTIONS STARTING WITH A PLAY SECTION (EITHER REAL OR SPECIAL) AND ENDING EITHER WITH END OF DATA OR START OF ANOTHER PLAY SECTION. ALTERNATIVELY ALL OF THE SECTIONS FOR ONE THEATRE ON ONE DATE.

SECTION ALL CHARACTERS BETWEEN ONE BLANK-ASTERISK-SMALL LETTER SEQUENCE AND THE NEXT SUCH SEQUENCE, EXCEPT THAT THE LAST SECTION IN A BLK ENDS AT THE END OF THE BLK. THE FIRST SECTION IN A BLK STARTS WITH THE FIRST BLANK-ASTERISK-SMALL LETTER SEQUENCE IN THE BLK. ALL CHARACTERS OF LOGICAL DATA PRECEDING THIS SEQUENCE ARE IGNORED, EXCEPT FOR ILLEGAL CHAR CHECKS. IN MTST INPUT THE '*' IS REPLACED BY THE PREFIX. IN SCANNER INPUT THE BLANK PRECEDING THE ASTERISK IS NOT ALWAYS REQUIRED, ALTHOUGH TO BE PROPER IT SHOULD ALWAYS BE THERE.

SECTION HEADER THE ITEMS FOLLOWING THE SECTION DELIMITER CHARACTER AND PRECEDING THE CAST LIST, IF ANY. FOR UNTITLED SECTIONS THIS IS THE SECTION TYPE LETTER. FOR TITLED SECTIONS THE HEADER STARTS WITH THE SECTION TYPE LETTER AND ENDS WITH THE FIRST OCCURRENCE OF ANY OF THE FOLLOWING ..

- 1) END OF THE SECTION.
- 2) PERIOD-BLANK-BLANK 3 CHAR SEQUENCE, '. '.
- 3) QUESTION MARK-BLANK-BLANK 3 CHARACTER SEQUENCE, '? '.
- 4) EXCLAMATION POINT-BLANK-BLANK 3 CHAR SEQUENCE, '! '.

SECTION TYPE LETTER (SECTYPE) THE CHARACTER FOLLOWING THE BLANK-ASTERISK OR BLANK-PREFIX DEFINING THE START OF A SECTION. THIS CHARACTER SHOULD BE A SMALL LETTER DESIGNATED TO REPRESENT A PARTICULAR TYPE OF SECTION.

PERFORMANCE DATE ENTRY IN A PLAY SECTION 1 TO 3 NUMBERS SEPARATED FROM EACH OTHER AND ALL OTHER ITEMS BY AT LEAST ONE BLANK. THE 1 TO 3 NUMBERS ARE CONSIDERED ONE ITEM, AND HENCE ANY PAGE ENTRY MUST PRECEDE ALL OF THEM. IF PRESENT THE DATE MUST LIE BETWEEN THE SECTION TYPE LETTER AND THE THEATRE, IF ANY (ABSENCE OF A THEATRE IN A PLAY SECTION IS AN ERROR).

THEATRE ENTRY A CHARACTER STRING IN THE HEADER PART OF A PERFORMANCE SECTION THAT FOLLOWS THE DATE ENTRY, IF ANY, SEPARATED FROM IT BY AT LEAST ONE BLANK, AND PRECEDES THE TITLE, IF ANY, SEPARATED FROM IT BY AT LEAST ONE BLANK. THE THEATRE BEGINS WITH THE FIRST NONBLANK CHARACTER, EXCEPTING PERIODS AND CERTAIN OTHER PUNCTUATION CHARACTERS, FOLLOWING THE DATE ENTRY, IF ANY. THE THEATRE ENDS WITH THE FIRST BLANK FOLLOWING ITS START, OR ELSE ENDS WITH THE END OF THE SECTION, WHICHEVER COMES FIRST. THE LEGAL CHARACTERS FOR A THEATRE ARE THE SMALL LETTERS AND THE APOSTROPHE.

TIME ENTRY TITLE TIME ENTRY OR CAST LIST TIME ENTRY.

TITLE AREA FOR A PLAY SECTION THE AREA FOLLOWING THE THEATRE ENTRY AND ENDING WITH THE END OF THE SECTION HEADER. FOR ANY OTHER TITLED SECTION THE AREA BETWEEN THE SECTYPE LETTER AND END OF SECTION HEADER.

FULL TITLE STRING A CHARACTER STRING IN THE TITLE AREA. IT STARTS WITH THE FIRST CHARACTER IN THE AREA THAT IS

MEMBER NAME DOCGENSP

NOT A PERIOD, BLANK, COMMA, SEMICOLON, QUESTION MARK, MINUS SIGN, OR EXCLAMATION POINT. IT ENDS WITH THE END OF THE TITLE AREA.

TITLE (SECTION TITLE) FOR PLAY AND AFTERPIECE SECTIONS THIS IS THE PART OF THE FULL TITLE STRING PRECEDING THE SUBTITLE IN THE STRING, IF ANY. FOR OTHER TITLED SECTIONS THIS IS THE PORTION OF THE FULL TITLE FOLLOWING THE TIME ENTRY, IF ANY, AND PRECEDING THE SUBTITLE, IF ANY. HOWEVER ONLY 120 CHARS ARE PUT OUT FROM THE STRUCTUR PROGRAM, AND ONLY 83 ARE AVAILABLE IN SORTS.

TITLE TIME ENTEY A CHARACTER STRING IN THE FULL TITLE OF A TITLED SECTION. IF PRESENT IT STARTS WITH THE FULL TITLE STRING AND ENDS WITH THE FIRST 2 CHARACTER SEQUENCE OF COLON-BLANK, ': ', IN THE FULL TITLE STRING. THE STRING ITSELF MUST BE A LEGITIMATE CAST LIST TIME ENTRY IN OTHER RESPECTS.

TRICK SECTION A TITLED SECTION TAGGED BY THE SMALL LETTER 'T'. ITS DATA SHOULD COME FROM TITLED ENTERTAINMENT BOXES.

TITLED SECTION A SECTION THAT IS SUPPOSED TO HAVE A TITLE. AT PRESENT THESE SECTIONS ARE TAGGED BY ..

A, P, I, O, T, U. DISCOUNTING PAGE ENTRIES, ALL THESE SECTIONS SHOULD HAVE A TITLE OF AT LEAST THE SPECIAL VARIETY.

STRUCTURED TEXT TEXT THAT IS NOT EXTRANEOUS TEXT.

STRUCTURED SECTION ANY KIND OF SECTION EXCEPT A COMMENT SECTION.

TITLE LADDER ENTRY A LADDER ENTRY IN EITHER A TITLED OR UNTITLED SECTION THAT IS NOT PRECEDED BY ANY CAST GROUP ITEM. .. THERE ARE TWO TYPES .. 'SEE' & 'AS'.

SYNTACTIC ROLE IS ANY ITEM THAT IS ..

- 1) IN A CAST GROUP THAT HAS A DASH OR CAST GROUP LADDER ENTRY IN IT.
- 2) TO THE LEFT OF ALL DASHES AND THE CAST GROUP LADDER ENTRY, IF ANY, IN ITS GROUP.
- 3) IS NOT A TIME ENTRY.
- 4) IS NOT A PAGE ENTRY.
- 5) IS NOT A LADDER ENTRY.

SYNTACTIC ACTOR IS ANY ITEM IN A CAST LIST GROUP THAT ..

- 1) IS TO THE RIGHT OF THE 1ST DASH IN THE GROUP IF THE GROUP HAS A DASH IN IT, OR IS TO THE RIGHT OF THE CAST GROUP LADDER ENTRY, IF ANY, IN A GRP THAT HAS NO DASH.
- 3) IS NOT A TIME ENTRY.
- 4) IS NOT A LADDER ENTRY.
- 5) IS NOT A PAGE ENTRY.

THE RIGHT HAND DELIMITER OF A SYNTACTIC ACTOR MAY BE ..

- 1) A SEMICOLON.
- 2) END OF GROUP (EITHER SEMICOLON OR END OF SECTION).
- 3) COMMA.

*** STANDARD DATA FILES ***

THERE ARE 11 BASIC KINDS OF FILES ENCOMPASSING 7 GENERAL DATA FORMATS IN USE BY THE LSP INPUT SYSTEM. THE 11 TYPES OF FILE ARE NAMED AND BRIEFLY DESCRIBED BELOW. MORE DETAILED DESCRIPTIONS ARE GIVEN IN THE INDIVIDUAL PROGRAM WRITEUPS.

LSPIC ORIGINAL DATA FILE OF SCANNER OUTPUT. LSPIC FILES ARE INPUTS TO THE ICISCAN, ICIFRONT, AND ICIFIX PROGRAMS, AND OUTPUT FROM THE ICIFIX PROGRAM.

LSPCR CORRECTION TEXT PORTION OF AN LSPIC FILE. FOR USE WITH

MEMBER NAME DOCGENSP
WITH ICIFRONT. ALSO PRODUCED BY THE SLASH PROGRAM FROM
CARD FORMAT INPUT.

LSPDT DATA TEXT PORTION OF AN LSPIC FILE. INPUT FOR
ICISCAN OR ICIFRONT OR ICIFIX. OUTPUT FROM ICIFIX OR
SCANNER (OR LSPCNLSL ETC.).

LSPMT ORIGINAL DATA FILE. LSPMT IS AN INPUT FOR THE SAVEDT PROGRAM,
IF IT IS DESIRED TO CREATE A BACKUP OF THE INPUT THEN
AN LSPMT FILE CAN ALSO BE OUTPUT FROM THE SAVEDT PROGRAM.
SIMILARLY LSPMT FILES ARE BOTH INPUT AND OUTPUT FOR EDT.

LSPCL STANDS FOR 'CLEANED'. THIS IS OPTIONAL INPUT/OUTPUT OF
THE ICISCAN PROGRAM. IT IS A LINE BY LINE FORM OF TEXT
THAT HAS BEEN PUT THROUGH THE LINE READIN PROCESS AND
CLEANED OF 'AT' SIGNS AND TYPIST ENTRIES ETC. EACH LINE
IS HEADED BY THE FOLLOWING PIECES OF INFORMATION IN A
FIXED FORMAT ..
1) TYPIST'S LETTER, 1 CHAR.
2) PAGE NUMBER, 4 CHAR.
3) LINE NUMBER WITHIN PAGE, 5 CHAR.
4) ABSOLUTE LINE NUMBER, 6 CHAR.
5) LENGTH OF LINE, 4 CHARS.
THUS A LSPCL FILE IS SUITABLE FOR USE AS A STREAM FILE
ALLOWING EASY MANIPULATION IN A VARIETY OF
WAYS BY A FORTRAN PROGRAM, EITHER BATCH OR RAX.

LSPPU STANDS FOR 'PURE'. THIS IS THE MAIN OUTPUT OF SAVEDT OR ICISCAN
AND INPUT TO THE STRUCTUR PROGRAM. IT IS TEXT CLEANED OF COMMENT
SECTIONS, AND EXTRANEOUS TEXT, AS WELL AS RUBOUT CHARS,
ILLEGAL CHARS, AND CERTAIN CONTROL TEXT. EACH BLK SHOULD BE
1 OR MORE COMPLETE SECTIONS, AND EACH BLK SHOULD CONTAIN
ONLY COMPLETE SECTIONS, NOT JUST PART OF A SECTION.

LSPST STANDS FOR STRUCTURED TEXT. IT IS OUTPUT FROM THE STRUCTUR
PROGRAM, AND INPUT TO THE LADDER PROGRAM. THE DATA IN THIS
FILE CONSISTS OF TAGGED GROUPS AND ITEMS, ONE SECTION PER BLK.
THIS IS A REGIONAL(3) U FORMAT KEYED SEQUENTIAL FILE.

LSPSS STANDS FOR 'STRUCTURED SEQUENTIAL', LIKE LSPST, EXCEPT
THAT IT HAS THE DIRECT ACCESS KEY OF THE LSPST FILE APPENDED
TO THE FRONT OF THE BLK, AND IT IS PADDED IN FRONT OF THE
KEY WITH 2 BLANKS TO MAKE A MINIMUM BLKSIZE OF 18 BYTES.
THIS FORMAT IS INTENDED FOR STRUCTURED OUTPUT TO TAPE.
THIS IS AN ORDINARY UNKEYED U FORMAT SEQUENTIAL FILE.

LSPLD STANDS FOR 'LADDER'. THIS IS OUTPUT FROM THE LADDER PROGRAM AND
AND INPUT TO THE ITEMGET PROGRAM. IT IS A KEYED DIRECT
ACCESS FILE HOLDING SECTIONS WITH LADDER
REFERENCES CONVERTED TO EXPLICIT CAST LISTS AND TITLES.
LSPLD FILES ARE VERY SIMILAR TO LSPST FILES.

LSPLS STANDS FOR 'LADDER SEQUENTIAL'. JUST LIKE LSPLD EXCEPT
THAT, AS IN THE LSPSS FILE, THE KEY IS APPENDED TO
THE FRONT OF THE BLK AND PRECEDED BY 2 BLANKS TO MAKE THE
18 BYTE MINIMUM BLKSIZE FOR TAPE. THIS IS AN ORDINARY
U FORMAT SEQUENTIAL FILE.

LSPIT STANDS FOR 'ITEMS'. IT IS OUTPUT FROM THE ITEMGET PROGRAM
AND INPUT TO THE SORT STEP. THE FORMAT IS FIXED
UNBLOCKED KEYED REGIONAL(3) DIRECT ACCESS,
AND EACH RECORD USUALLY CORRESPONDS TO A SYNTACTIC ROLE/ACTOR
PAIR. HOWEVER THE RECORD MAY HOLD AS LITTLE AS THE SECTION
TYPE, DATE, AND THEATRE.

LSPIS STANDS FOR 'ITEMS SEQUENTIAL'. SIMILAR TO LSPIT, BUT
INTENDED FOR UNKEYED SEQUENTIAL TAPE OUTPUT FROM THE
ITEMGET PROGRAM. THIS IS AN UNBLOCKED F FORMAT FILE.

MEMBER NAME DOCGENSP

LSPSR STANDS FOR 'SORTED'. OUTPUT FROM THE SORT PROGRAM.

SAME FORMAT AS LSPIS, EXCEPT THAT MAY BE BLOCKED, IF BLOCKED, IS USUALLY BY A FACTOR OF 19. INPUT FOR THE FORMAT PROGRAM.

LSPIX SIU INDEX ENTRIES. OUTPUT FROM THE SAVEDT PGM. AT PRESENT ONLY CONTAINS THE INDEX ENTRY ITSELF. SHOULD ALSO CONTAIN THE PAGE NUMBER, DATE, AND TAG. ICISCAN CAN EASILY BE ALTERED TO PUT OUT LEGITIMATE SIU INDEX ENTRIES ENCOUNTERED IN APPROPRIATE CONTEXTS.

1

CURRENT STATUS OF PROGRAMMING

PROGRAMMING TO BE DONE ..

- 1) INDIRECT SORT SYSTEM.
- 2) OPTIONAL CORRECTION OF ROLES OR ACTORS THAT START WITH 'END' ETC.
- 3) IMPLEMENT QUESTION MARK AS TITLE DELIMITER FOR TITLED SECTIONS.
- 4) ALTER LADDER TO PREVENT CGLE UPDATES FROM HAVING QUESTION MARKS APPENDED TO THEM.
- 5) MINOR IMPROVEMENTS IN ICIFRONT.
- 6) PUT RECORD COUNT IN FRMAT, OPTIONAL PRINT OUTPUT, AND OPTIONAL CHECK FOR IDENTICAL RECORDS.

STATUS OF PROGRAMS ..

CARDIN	DISCONTINUED
EDT	DONE
SAVEDT	DONE
ICIFIX	DONE
ICIFRONT	MAINT STAGE
ICISCAN	MAINT STAGE
STRUCTUR	MAINT STAGE
LADDER	MAINT STAGE
ITEMGET	MAINT STAGE
FORMAT	MAINT STAGE
LSPRINT	DONE
DOCPRINT	DONE
PLIST	DONE
TRT	DONE
MAP	DONE
COPYCAT	DONE
ICIORDER	NOT DONE
SLASH	MAINT STAGE

DOCUMENTATION ..

PROGRAMMING SPECS	MAINT STAGE
INPUT SPECS	MAINT STAGE
GENERAL	HALF DONE
BUG NOTES	NOT DONE
CHARACTER SET TABLES	DONE
OPERATION INSTRUCTIONS	MAINT STAGE
LISTINGS	DONE
RECOMMENDATIONS	NOT DONE

PROCEDURES ..

LSPCNLSL	DONE
LSPICFIX	DONE
LSPSLASH	NOT DONE
LSPFRONT	DONE
LSPSCLIN	DONE

MEMBER NAME	DOCGENSP
LSPRINTS	DONE
LSPISCAN	DONE
LSPSTRCT	DONE
LSPLADDR	DONE
LSPITEMS	DONE
LSPFRMAT	DONE
LSPSORT4	DONE
LSPSORTT	DONE
LSPMERGE	DONE
LSPBACMP	DONE
LSPUTGEN	DONE
LSPUPGEN	DONE
LSPROCUP	DONE

MEMBER NAME DOCSFORM

**** ORIGINAL MAGTAPE INPUT FROM SCANNER ****

THE SCANNER COMPANY PRODUCES NONLABELLED TAPE FORMATTED THUS ..
DCB=(RECFM=FB,BLKSIZE=800,LRECL=80).

EACH LOGICAL RECORD NOMINALLY CORRESPONDS TO A LINE OF TEXT
ON THE CHINA DATA SHEETS. THERE MAY BE AN OCCASIONAL EXTRA
LINE OF 80 BLANKS, USUALLY OCCURRING AT THE END OF A PAGE.
EACH PAGE IS SUPPOSED TO HAVE 30 LINES. EACH BLK OF TEN 80 BYTE
RECORDS IS ONE THIRD OF A PAGE. IF THERE IS AN EXTRA BLANK LINE
THEN THE OTHER LINES FOLLOWING IT ARE PUSHED DOWN EVEN INTO THE
NEXT BLK IF NECESSARY TO KEEP THE BLKSIZE AT 800.

THE TYPED PORTION OF EACH LINE IS NOMINALLY THE FIRST 75
CHARS OF THE LINE, THE REMAINING 5 CHARS BEING BLANK ON THE CHINA
DATA SHEETS EXCEPT IN CASE OF ERROR. AFTER SCANNING TO TAPE
THIS IS STILL GENERALLY TRUE, BUT OFTEN A TYPED DOUBLE QUOTE CHAR
IS SCANNED AS 2 APOSTROPHES, THUS CAUSING THE 75TH TYPED CHAR TO
SPILL OVER INTO COLUMN 76. TO SOLVE THIS PROBLEM ANY PAIR OF
CONTIGUOUS APOSTROPHES FOUND ON A LINE IS REPLACED WITH A
DOUBLE QUOTE CHAR BY THE ICISCAN PROGRAM. OCCASIONALLY NONBLANK
CHARS ARE PUT INTO COLUMN 76 OR BEYOND FOR OTHER REASONS, ALL
ERRORS OF SOME SORT. THESE ERRORS ARE ACCEPTED UP TO THE LAST
NONBLANK CHAR ANYWAY. NO ERROR MESSAGE IS GENERATED FOR THIS.

THE CHARACTERS ARE STANDARD EBCDIC EXCEPT FOR THE FOLLOWING ..

THE BRITISH POUND SIGN IS AN EBCDIC NUMBER SIGN, HEX '7B'.
IT IS PRINTED BY THE LSPRINT SN TABLES AS AN 'L' OVERPRINTED
ON A MINUS SIGN. IN THE ICIFIX PROGRAM THE 1052 PRINTS IT AS JUST WHAT
IT IS INTERNALLY, THE NUMBER SIGN. NEITHER THE HN NOR THE SN
CHAIN CAN PRINT THE NUMBER SIGN CHARACTER.

THE LEFT AND RIGHT BRACKETS ARE HEX 'AD' AND 'BD' RESPECTIVELY.
THEY PRINT AS THE PARENTHESES OVERPRINTED WITH THE MINUS SIGN
USING THE LSPRINT SN TABLES. ON THE 1052 THEY ARE PRINTED AS
THE 'LESS THAN' AND 'GREATER THAN' SYMBOLS RESPECTIVELY BY ICIFIX.

THE CHARACTER THE SCANNER COMPANY IS SUPPOSED TO USE FOR
NONRECS IS THE HEX '00'. THIS IS PRINTED BY THE SN LSPRINT TABLES
AS A ZERO OVERPRINTED ON THE PREFIX CHAR. OTHER ILLEGAL CHARS
PRINT AS THE USUAL 'H' ON 'I' CHAR.

MEMBER NAME DOCOLSPS
OPERATING THE LONDON STAGE PROJECT INPUT SYSTEM

SUMMARY OF STEPS

*** INITIAL PROCESSING ***

LSPBAKUP TO CREATE SL BACKUP OF NL TAPE FROM SCANNER COMPANY.
ICIFIX TO PRODUCE A FILE OF DATA TEXT FROM THE SL BACKUP,
DELETE BLANK LINES, AND LOCATE BAD TYPIST ENTRIES.
ICIFIX TO PRODUCE A FILE OF CORRECTION TEXT FROM THE SL,
BACKUP, DELETING BLANK LINES.
LSPRINT TO LIST THE FILE OF CORRECTION TEXT.
ICISCAN TO PRINT ONTO TAPE A LIST/LOG OF THE DATA TEXT AND TO
PRODUCE PURE FILE OUTPUT FOR THE STRUCTUR PROGRAM.
LSPRINT TO LIST ICISCAN'S LIST/LOG PRINTOUT WITH OVERPRINTING
ON THE SN CHAIN FOR ERROR ANALYSIS AND CROSS REFERENCE.
STRCT FOR LOCATING DCHG ERRORS AND SEASON CHANGES THAT
DON'T WORK RIGHT.
LADDR TO LOCATE ANY SINGLE ERRORS THAT CAUSE MANY ERROR
MESSAGES, BUT ARE EASY TO FIX. THIS STEP SHOULD
NOT BE RUN UNLESS THE STRCT STEP HAS NO BAD DCHG ERRORS,
AND NO SEASON CHANGES MISSED.

*** FIRST CYCLE ***

ANALYSIS (DONE BY THE OPERATOR) TO LOCATE IMPORTANT ERRORS,
AND DETERMINE REMEDIAL ACTION FOR THEM.
THE FOLLOWING SPECIAL AREAS FOR THE OPERATOR TO
CHECK ARE LISTED IN ORDER OF IMPORTANCE ..
1) THE FIRST LINE.
2) THE ENDING OF EACH SEASON EXCEPT THE LAST IN THE RUN.
3) THE START OF EACH MONTH OR YEAR.
4) TYPIST ENTRIES.
5) SEASONS NOT STARTING WITH A PLAY SECTION. IN THIS
EVENT THE ITEMGET PROGRAM MAY BOMB OUT, BUT LADDR
SHOULD WORK PRETTY WELL.
ICIFIX TO FIX SERIOUS ERRORS, AND POSSIBLY OTHER ERRORS,
DISCOVERED BY ICIFIX, ICISCAN, STRCT, AND POSSIBLY LADDR.
ICIFRONT (AN OPTIONAL STEP) TO APPLY CORRECTIONS TO THE DATA TEXT.
ICISCAN (EITHER USING LSPSCLIN FOR CLEANED FORMAT INPUT OR
LSPISCAN FOR SCANNER FORMAT DATA TEXT) TO CHECK FOR
SERIOUS ERRORS, ESPECIALLY FROM APPLICATION OF CORRECTIONS.
OPTION 6 SHOULD BE USED INSTEAD OF OPTION 3.
STRCT SAME AS FOR INITIAL PROCESSING STAGE.
LADDR SAME AS FOR INITIAL PROCESSING STAGE.

*** SECOND (& LATER) CYCLES ***

ANALYSIS AS IN FIRST CYCLE, BUT MAY INVOLVE DEVELOPMENT OF FIXES
ON CARDS FOR CLEANED FORMAT OUTPUT FROM ICIFRONT, TO
BE CORRECTED AGAIN WITH ICIFRONT.
ICIFIX FOR DATA TEXT. AS IN FIRST CYCLE, BUT MAY BE SUPERSEDED
BY USE OF SLASH PROGRAM.
ICIFIX FOR CORRECTION TEXT. MAY BE SUPERSEDED BY SLASH PROGRAM.
SLASH MAY BE USED TO CONVERT CORRECTION TEXT ON CARDS TO
PROPER FORMAT INPUT TO ICIFRONT'S CORRECTION TEXT FILE.
ICIFRONT MAY BE USED (OR NOT) EITHER WITH SLASH OUTPUT OR
ORIGINAL CORRECTION TEXT, EITHER WITH CLEANED FORMAT
INPUT OR SCANNER FORMAT INPUT (4 COMBINATIONS).

MEMBER NAME DOCOLSPS
ICISCAN SAME AS FIRST CYCLE.
STRCT SAME AS OTHER RUNS.
LADDR SAME AS OTHER RUNS.

*** RETRIEVAL STEPS ***

ITEMS TO SELECT SOME OR ALL OF SORT RECORDS REPRESENTED BY LADDR OUTPUT.
SORT TO SORT SORT RECORDS IF NECESSARY.
MERGE (ACTUALLY PART OF SORTING PROCESS) TO COMBINE OUTPUT OF SEVERAL SORT RUNS IF NECESSARY.
FRMAT TO LIST SORT RECORDS.

**** INITIAL PROCESSING SEQUENCE ****

WHEN THE ORIGINAL MAGTAPE ARRIVES FROM THE SCANNER COMPANY THE FIRST THING TO DO IS TO MAKE A BACKUP COPY ON A STANDARD LABELLED TAPE USING THE LSPCNLSL PROCEDURE. BE SURE TO TAKE THE WRITE RING OFF THE SCANNER COMPANY'S TAPE BEFORE USING IT, FOR SAFETY.

THREE PARAMETERS MAY HAVE TO BE SPECIFIED ON THE EXEC CARD IN THE USUAL CASE ..

- 1) OUTVOL= IS THE VOLUME SERIAL NUMBER OF THE STANDARD LABELLED OUTPUT TAPE. DEFAULT IS 'A'.
- 2) OUTSEQ= IS THE OUTPUT FILE SEQUENCE NUMBER. DEFAULT IS 1.
- 3) OUTDSN= IS THE DATASET NAME OF THE OUTPUT FILE. THE DEFAULT IS 'LSPIC'.

LSPIC FILE NAMES ARE DESIGNED TO DESCRIBE THE INFORMATION IN THE FILE. THEY ARE 8 CHARACTERS LONG. EACH CHARACTER HAS THE FOLLOWING MEANING ..

- (1) THE YEAR IN WHICH THE DATA IS STORED.
A=71, B=72 ETC.
- (2) THE MONTH IN THIS YEAR IN HEX, 5 = MAY, C = DEC., ETC.
- (3) THE DAY OF THE MONTH BY THE FOLLOWING SYSTEM .. 1-9 = 1-9,
A=10, B=11, ETC.
- (4-6) THE LAST THREE DIGITS OF THE YEAR IN WHICH THE DATA IN THE FILE STARTS.
- (7) THE MONTH IN WHICH THE DATA STARTS, USING THE SAME CODES AS AS FOR CHARACTER (2) ABOVE.
- (8) THE DAY ON WHICH THE DATA STARTS, SAME CODES AS (3) ABOVE.

OF COURSE USE OF THIS NAMING CONVENTION IS NOT REQUIRED, IT IS ONLY INTENDED FOR SAFETY AND CONVENIENCE IN KEEPING TRACK OF TAPES. IF THE USER FINDS IT INCONVENIENT THEN IT SHOULD BE IGNORED.

THE SCANNER COMPANY'S TAPE SHOULD BE MOUNTED ON UNIT 182, AND THE OUTPUT TAPE ON UNIT 180. THESE CAN BE ALTERED BY USING INUNIT= AND OUTUNIT= ON THE EXEC CARD. EXAMPLE

//JJJK EXEC LSPCNLSL,OUTVOL=LSP009,OUTDSN=A4B777C6,INUNIT=181

HERE THE INPUT UNIT IS 181 INSTEAD OF THE DEFAULT 182, AND THE OUTSEQ= PARAMETER WAS NOT NEEDED SINCE THE FILE SEQUENCE WAS TO BE 1.

AFTER THE BACKUP STEP IS COMPLETED REMOVE THE SCANNER COMPANY'S TAPE. KEEP IT UNTIL 2 OTHER COPIES OF ITS DATA HAVE BEEN MADE AND CHECKED.

AFTER THE BACKUP, THE NEXT STEP IS TO SEPARATE THE CORRECTION TEXT FROM THE DATA TEXT, CREATING TWO NEW FILES WITH THE ICIFIX PROGRAM. WHEN CREATING THE DATA TEXT ENABLE TYPIST ENTRY CHECKING AND LEAVE A RECORD OF BAD PAGE ENTRIES ON THE CONSOLE PRINTOUT. AT THIS POINT

MEMBER NAME DOCOLSPS

DO NOT ALTER THE TEXT IN ANY WAY EXCEPT FOR SEPARATING THE TWO KINDS OF TEXT AND POSSIBLY DELETING BLOCK FILLER AT THE END OF THE FILE. THE REASON FOR NOT TAMPERING IS TO PRESERVE THE ORIGINAL DATA TEXT FOR ACCOUNTING PURPOSES.

ICIFIX OPERATION IS DESCRIBED IN THE ICIFIX PROGRAM WRITEUP. WHEN A REVISION IS MADE OF A BATCH THEN THE DATE OF STORAGE, GIVEN IN COLUMNS 1-3 OF THE NEW DSN, SHOULD REFLECT THE DATE OF THE REVISION.

AFTER SEPARATION PRINT THE CORRECTION TEXT USING THE LSPRINTS PROCEDURE WITH OPT=DIRECTSN. IF YOU ARE ANXIOUS TO SEE THE ORIGINAL DATA TEXT THEN, OF COURSE, GO RIGHT TO THE ICISCAN LIST/LOG STEP.

THE VERY FIRST ICISCAN RUN, THE LIST/LOG STEP, IS DONE USING THE LSPISCAN PROCEDURE, WHOSE COMMONLY REQUIRED PARAMETERS ARE ..

- 1) INVOL INPUT VOLUME LABEL.
- 2) INDSN INPUT DATASET NAME.
- 3) INSEQ INPUT FILE SEQUENCE NUMBER, DEFAULT IS 1.
- 4) OUTSEQ OUTPUT FILE SEQUENCE NUMBER, DEFAULT IS 1.

THE RUN DECK SHOULD HAVE 2 DD CARDS FOR SYSPRINT, 1 FOR THE PRINTER DIRECTLY, THE OTHER FOR OUTPUT TO LSP002 ON UNIT 181. PLACE THE ONE THAT YOU WANT TO BE USED AHEAD OF THE OTHER ONE IN THE DECK. IF YOU ARE USING TAPE FOR SYSPRINT THEN IT IS SOMEWHAT MORE EFFICIENT TO PLACE THE LSPRINT STEP IMMEDIATELY AFTER THE ICISCAN STEP AND HAVE THE TAPE PRINTED IMMEDIATELY.

IN GENERAL A TAPE SHOULD BE USED FOR SYSPRINT ON THE VERY FIRST RUN. THIS SPEEDS UP PROCESSING AND MAKES POUND SIGNS, BRACKETS, AND EQUAL SIGNS VISIBLE WHEN LSPRINT PRINTS THE TAPE. THIS WILL ALSO HELP LOCATE ANY ILLEGAL CHARACTERS THAT MAY BE PRESENT. THE PRINT FILE LISTS, IN ADDITION TO THE RAW INPUT TEXT, ALL ERROR MESSAGES, THE ABSOLUTE LINE NUMBER OF EACH LINE, AN APPROXIMATE APPARENT DATE ASSOCIATED WITH EACH LINE, AND THE APPROXIMATE STRCT/LADDR/ITEMS INPUT BLK NUMBER ASSOCIATED WITH EACH LINE. THE TYPED PAGE NUMBER AND LINE NUMBER CAN EASILY BE DETERMINED BY INSPECTION OF THE TEXT. IN SHORT THIS LISTING PROVIDES A CROSS-REFERENCE TOOL CLOSELY CONNECTING SUCH DIVERSE ENTITIES AS THE ORIGINAL DATA IN THE LONDON STAGE ITSELF AND THE FINAL SORTED FORMATTED OUTPUT. FORMATTED OUTPUT HAS GONE THROUGH AT LEAST 6 MAJOR TRANSFORMATIONS SINCE ITS 'ORIGIN' IN THE LONDON STAGE.

PRINTING OF ICISCAN'S LIST/LOG SHOULD BE DONE WITH THE LSPRINTS PROCEDURE USING OPT=PRINTSN.

IT IS ALRIGHT IF THE PRINTER IS ALLOWED TO RUN OUT OF PAPER IN THIS STEP. THE OUTPUT WILL BE BULKY ANYWAY, BUT THE NUMBER OF PIECES SHOULD NOT EXCEED 3 IF POSSIBLE. THE PRINTOUT IS INTENDED MAINLY FOR REFERENCE IN THE ERROR CORRECTION PROCESS.

IF THE QUALITY IS NOT TOO LOW THEN, AFTER CORRECTION NOTES ETC. HAVE BEEN MADE ON IT, THIS PRINTOUT SHOULD BE SENT TO DR. SCHNEIDER. SCRAP PAPER SHOULD BE USED FOR THIS PRINTOUT IF THERE ARE ANY SIZEABLE CHUNKS AVAILABLE.

HERE IS A LIST OF ITEMS REQUIRING SPECIAL ATTENTION IN RUNNING STRCT, ESPECIALLY FOR THE FIRST TIME ON A NEW BATCH ..

- 1) CHECK THE MONTH ABBREVIATIONS USED IN LADDER REFERENCES, AND CHANGE THE MONSTR CARDS ACCORDINGLY. SOME PARTS USE 'APR' WHILE OTHERS USE 'APRIL', ETC. (MARCH, APRIL, MAY, JUNE, JULY, AND AUGUST ARE THE ONES TO WATCH).

MEMBER NAME DOCOLSPS

2) IT IS IMPORTANT TO BE SURE ABOUT THE SEASON ENDS IF YOU HOPE TO RUN THE OUTPUT THROUGH LADDR. ONE LIKELY ERROR IS COUNTING SECTIONS IN THE LAST PERFORMANCE OF A SEASON THAT ARE NOT PUT OUT BY ICISCAN. ANOTHER IS OVERLOOKING THE DELETION OF AN ASTERISK BY AT SIGNS. ILLEGAL SECTIONS THAT HAVE CAPITAL SECTYPE LETTERS ARE PUT OUT BY ICISCAN AS REGULAR SECTIONS WITH LEGAL SMALL LETTERS AS LONG AS THE SMALL LETTER IS THAT OF A LEGAL SECTYPE THAT HAS BEEN SELECTED ON THE OPTION CARD.

3) CHECK THE JCL CURSORILY TO MAKE SURE NO UNWANTED OVERRIDE CARDS ARE THERE. STRCT AND LADDR ARE SOMEWHAT DANGEROUS IN THIS RESPECT SINCE THEY ARE EXECUTED BY STRAIGHT JCL, NOT PROCEDURES. SPECIAL CARDS LEFT OVER FROM THE PREVIOUS RUN ARE LIKELY TO CAUSE TROUBLE.

THE LADDER STEP SHOULD NOT BE RUN UNTIL THE STRUCTURE STEP PRODUCES NO BAD 'DCHG' MESSAGES. THIS IS ONE IN WHICH THE DATE IS OFF BY 5 DAYS OR MORE. ADDITIONALLY ALL SEASON CHANGES MUST OCCUR PROPERLY. A MERE CHANGE IS NOT ENOUGH, IT MUST BE ENTIRELY CORRECT.

*** ERROR CORRECTION CYCLES ***

AFTER THE FIRST CYCLE THE ERROR CORRECTION LOOP ANALYSIS STEP MAY INCLUDE ANALYSIS OF CORRECTION TEXT ERRORS AND A COMPARISON OF ICISCAN AND/OR STRCT ERRORS IN USING ICIFRONT VS. NOT USING IT. THE MOST LIKELY SERIOUS ERROR IN CORRECTION TEXT INVOLVES A CORRECTION SPECIFICATION THAT INDICATES AN INCORRECT TYPIST OR ADVANCES THE PAGE NUMBER TOO MUCH. THIS CAN USUALLY BE SPOTTED WHEN THERE ARE VERY MANY 'CORRECTIONS' DONE TO ONE LINE IN THE UPDATE LOG PRINTOUT, OR WHEN THE PROGRAM STOPS AND THERE ARE STILL CORRECTIONS LEFT TO DO.

THE ICIFIX DATA TEXT STEP IS AS DESCRIBED FOR THE INITIAL PROCESSING SEQUENCE, EXCEPT THAT THE INPUT IS THE MOST RECENT VERSION OF THE DATA TEXT.

IF ICIFRONT IS TO BE USED AGAIN AND ERRORS HAVE BEEN FOUND PREVIOUSLY THEN THE ICIFIX CORRECTION TEXT CORRECTION TEXT STEP WILL HAVE TO BE DONE. DO NOT ENABLE TYPIST ENTRY CHECKING FOR THIS STEP. THIS MUST BE DONE BY HAND FOR CORRECTION TEXT.

USE THE LSPFRONT PROCEDURE TO RUN ICIFRONT. THE STANDARD UNITS ARE .. 181 FOR DATA TEXT, 182 FOR CORRECTION TEXT, AND 180 FOR CLEANED FORMAT OUTPUT. THE RESULTS MAY LOOK FAIRLY BAD AT FIRST, BUT INITIALLY ICIFRONT'S OUTPUT IMPROVES DRAMATICALLY WITH JUST A SMALL AMOUNT OF FIXING IN BOTH THE DATA AND CORRECTION TEXTS.

IF ICIFRONT WAS NOT USED THEN THE ICISCAN STEP SHOULD BE RUN WITH THE LSPISCAN PROCEDURE AS IN THE INITIAL PROCESSING STAGE. THE ICIFRONT CASE IS DESCRIBED IN THE FOLLOWING PARAGRAPHS ..

LSPSCLIN IS USED TO RUN THE CLEANED FORMAT OUTPUT THROUGH THE ICISCAN PROGRAM. THE INPUT FILE IS NAMED CLIN. THE DEFAULT INPUT UNIT IS 180 AS WITH THE LSPISCAN PROCEDURE. THE STANDARD OUTPUT UNIT AND VOLUME ARE ALSO IDENTICAL TO THOSE OF THE LSPISCAN RUN. TO AVOID OVERWRITING PREVIOUS OUTPUT YOU MAY USE A HIGHER OUTSEQ NUMBER (AS LONG AS THERE IS ROOM ON THE TAPE).

ICIFRONT SEEMS TO INCREASE THE ERROR RATE WHEN USED ON RAW DATA AND CORRECTION TEXT. HOWEVER WHEN THE DATA TEXT HAS HAD ALL OF ITS BAD TYPIST ENTRIES FIXED AND CERTAIN HORRIBLE CORRECTION TEXT

MEMBER NAME DOCOLSPS

ERRORS HAVE BEEN FIXED THEN USE OF ICIFRONT IS LIKELY TO BE BENEFICIAL. THE REAL STICKLER IS THAT IT MAY CREATE A NEW BAD DCHG ERROR. ANOTHER PROBLEM IS THAT AN ERROR MAY BE CORRECTED BY BOTH ICIFRONT AND ICIFIX. THE SOLUTION HERE IS TO DELETE THE CORRECTION FROM THE CORRECTION TEXT AND CORRECT THE DATA TEXT WITH ICIFIX.

ALTHOUGH NOT ESSENTIAL, IT IS PROBABLY BETTER TO PUT THE SYSPRINT OUTPUT ON TAPE AND USE LSPIRINT RATHER THAN PRINT DIRECTLY, SINCE BRACKETS ARE SO RELEVANT AT THIS STAGE.

AFTER RUNNING ICISCAN A DECISION MUST BE MADE AS TO WHETHER TO GO ON WITH THE PRESENT CYCLE OR TO GO BACK AND USE ICIFIX TO CORRECT MAJOR ERRORS AND THEN RERUN ICISCAN AGAIN. THIS DECISION SHOULD BE BASED PRIMARILY ON THE SAME DETAILS AS ARE DESCRIBED IN THE ANALYSIS STEP. USUALLY IT IS BEST TO GO AHEAD AND RUN STRCT REGARDLESS OF ERRORS.

IN RUNNING STRUCTUR IT IS PROBABLY BEST TO JUST PRINT DIRECTLY ON THE PRINTER RATHER THAN USE UNIT 181 UNLESS IT IS ONE OF THE FINAL CYCLES OF THE DATA ENTRY STAGE. THERE IS NO TAPE PROCEDURE FOR THIS STEP, A STRAIGHT PROGRAM IS USED SINCE IT IS ANTICIPATED THAT NO FLEXIBILITY WILL BE NEEDED IN TAPE ALLOCATION ETC.

IN RUNNING STRCT THE SAME CONSIDERATIONS FOR OVERWRITING PREVIOUS OUTPUT APPLY AS WITH THE ICISCAN PROGRAM. SINCE STRCT IS RUN BY STRAIGHT JCL THE ACTUAL DD CARD IN THE RUN DECK WILL HAVE TO BE REPLACED OR SUPERSEDED BY ONE WITH THE SAME NAME IN THE DECK.

CONDITIONS FOR CONTINUING WITH LADDR RATHER THAN RESTARTING AT THE ANALYSIS STEP ARE ALWAYS THE SAME - NO BAD DCHG ERRORS, AND NO SEASON CHANGES MISSED.

*** RETRIEVAL STEPS ***

IF SORT RECORDS ARE TO BE PRODUCED THEN AFTER A SUCCESSFUL LADDR RUN YOU SHOULD ALMOST CERTAINLY RUN ITEMS RATHER THAN RESTART THE CYCLE. ONLY RESTART IF THERE IS AN EASY WAY TO GET RID OF A SUBSTANTIAL PORTION OF ERRORS.

TO SELECT ALL RECORDS IN A BATCH USE THE FOLLOWING STATEMENTS AFTER THE OPTION CARD ..

- 1) DATE RANGE, '5 1659 07 01 1800 12 31'
- 2) SE CARD TO CAUSE SELECTION OF ALL RECORDS, '1T'
- 3) GO CARD, '7'
- 4) STOP CARD, '9'

ITEMS SHOULD TAKE APPROXIMATELY 40 MINUTES PER OUTPUT REEL, PLUS SEVERAL MINUTES CHANGEOVER. BATCH 1 TOOK APPROXIMATELY 2 REELS. BEFORE RUNNING CHECK AND SEE IF ANY BATCH (ESP NO. 3) HAS A MUCH BIGGER OUTPUT CHAR COUNT IN THE LADDR PROGRAM, AND IF SO THEN ADD MORE NAMES TO THE OUTPUT VOLUME LIST ON THE ITSM DD STATEMENT. THE STEPNAME IS 'GO'.

THE SORT STEP, IF ANY, NEED NOT USE ANYTHING PECULIAR TO THE LONDON STAGE PROJECT SINCE THE PROGRAM INVOLVED IS SUPPLIED BY IBM, HOWEVER THERE ARE 2 PROCEDURES PROVIDED, LSPOSRTT AND LSPMERGE. LSPOSRTT IS USED FOR SORTS HAVING TAPE INPUT AND OUTPUT, USING DISKS AS INTERMEDIATE STORAGE AREAS. WITH SUFFICIENT SPACE ON THE DISKS IT CAN SORT APPRX 50,000 RECORDS PER RUN. THIS IS MORE THAN 1 TAPEFULL (UNBLOCKED) BUT LESS THAN 2 TAPES FULL. BATCH NUMBER 1 (WITH ONLY SLIGHT CORRECTIONS IN IT) HAS APPROXIMATELY 64,000 SORT RECORDS. A FAIRLY GOOD WAY

MEMBER NAME DOCOLSPS

TO GET AROUND THIS IS TO 1) DO N SEPARATE SORTS, EACH PRODUCING ONE TAPE OF OUTPUT FROM ONE TAPE OF INPUT, AND 2) THEN DO SUCCESSIVE 2 WAY MERGES OF THESE UNTIL ONLY ONE FILE (OF 1 OR MORE TAPES) IS LEFT. THIS MERGING PROCESS IS FACILITATED BY THE LSPMERGE PROCEDURE.

THE OUTPUT OF THE SORT STEP CAN BE USED AS INPUT TO ANOTHER SORT STEP, BUT IF OUTPUT IS BLOCKED THEN ONE SORT STEP MAY NOT BE ABLE TO HANDLE AN ENTIRE TAPE AT ONE SHOT. ONE SOLUTION IS TO USE THE LSPDBLOK PROCEDURE TO UNBLOCK THE DATA ONTO SEVERAL TAPES. AN UNBLOCKED TAPE CAN HOLD APPX. 32,000 RECORDS. A TAPE BLOCKED BY A FACTOR OF 19 CAN HOLD APPX. 96,000 RECORDS.

THE FRMAT PROGRAM PRINTS, IN A HOPEFULLY NEAT MANNER, OUTPUT FROM THE ITEMS OR SORT STEPS. THE TAPE PROCEDURE FOR THIS IS LSPFRMAT. THE NORMALLY REQUIRED PARAMETERS ARE INUNIT=, & INVOL=. INVOL= MAY ACTUALLY BE A LIST OF SEVERAL VOLUMES ENCLOSED IN QUOTES, FOR EXAMPLE ..

//F EXEC LSPFRMAT,INUNIT=180,INVOL='LSP018,LSP034'
DEFAULTS ARE INUNIT=182,INVOL=LSP018.

MEMBER NAME DOCOPERL

THE GROOVE

FOR STANDARDIZATION AND SIMPLICITY OF OPERATION, IT IS RECOMMENDED
THAT THE FOLLOWING UNITS, VOLUMES, AND DSN'S BE USED . .

INPUT	OUTPUT
-------	--------

PROCNAME	UNIT	VOLUME	DSNAME	UNIT	VOLUME	DSNAME
IEHINITT	ANY	VARIES	(NONE)			
LSPBAKUP	NL	182	-----	SL	180	VARIES VARIES
LSPICFIX	SL	182	VARIES	VARIES	SL	180
LSPFRONT	SL	181	VARIES	LSPDT		
"	SL	182	VARIES	LSPCR	SL	180
LSPISCAN	SL	180	VARIES	VARIES	SL	182
LSPSCLIN	SL	180	VARIES	LSPCL	SL	182
LSPSTRCT	SL	182	LSP010	LSPPU	SL	180
LSPLADDR	SL	180	LSP011	STRS	SL	182
LSPITEMS	SL	182	LSP012	LSPLD	SL	180
LSPSORTT	SL	180	LSP013	LSPIT	SL	182
LSPMERGE	SL	ANY	VARIES	LSPSR		
"	SL	ANY	VARIES	LSPSR	SL	ANY
LSPFRMAT	SL	182	LSP014	LSPSR	PRINTER	
LSPRINTS	SL	181	LSP002	SYSOUT	PRINTER	

OPTION BITS FOR ICIFRONT .. (IOPT)

- 1) CLEAN INPUT FROM FILE DATX INSTEAD OF SCANNER FORMAT INPUT.
 - 2) ENABLE LOG MESSAGES.
 - 3) ENABLE 'ORG FOR C/D' LOG MESSAGES.

EVERY COLUMN OF THE OPTION CARD SHOULD HAVE A '0' IN IT
UNLESS IT IS INDICATING A '1' OPTION.

OPTION BITS FOR ICISCAN

- 1) SIMULATE END OF DATA INPUT IMMEDIATELY (STOP).
 - 2) PRINT ALL STRUCTURED OUTPUT SECTIONS ALONG WITH THEIR LINE NUMBERS.
 - 3) LIST INPUT LINES AND CROSS REFERENCE INFO ON SYSPRINT.
 - 4) SUPPRESS 'BAD TYPIST ENTRY' MESSAGE FOR ENTRIES ENDING WITH '\$@' OF THE USUAL VARIETY.
 - 5) LIST SECTIONS THAT ARE PROBABLY IN ERROR DUE TO THEIR ASSOCIATION WITH AN ILLEGAL DELIM OR THEIR NOT HAVING A SMALL LETTER AFTER THE '*'.
 - 6) NOTE TYPIST PAGE ENTRIES AS THEY ARE ENCOUNTERED IN LINE READIN.
 - 7) CLEANED INPUT FROM FILE CLIN.
 - 8) PRINT PIECE OF TEXT PRECEDING ILLEGAL DELIM, FROM DELIM BEFORE ILLEGAL ONE TO THE ILLEGAL ONE INCLUSIVE, IMMEDIATELY FOLLOWING THE ILLEGAL DELIM MESSAGE ITSELF.
 - 9) PRINT A LINE OF CERTAIN DEBUGGING INFO ON ENCOUNTERING A DELIM.
 - 11) PERFORM ILLEGAL CHAR CHECK.
 - 12) ENABLE ELLEGAL DELIM ERROR MESSAGE.
 - 13) PRINT INDEX ENTRIES LEGITIMATELY PROCESSED. CERTAIN INCORRECTLY DELIMITED ONES OR OUT OF CONTEXT ONES ARE NOT CAUGHT.
 - 14) CLEANED OUTPUT TO FILE CLOUT.
 - 15) SKIP INPUT LINES, BUT STILL COUNT THEM AS USUAL - IN PARTICULAR BLANK LINES ARE STILL NOT COUNTED.
 - 16) REVERT TO REGULAR TEXT MODE IF IN CORRECTION TEXT MODE ONLY UPON ENCOUNTERING A PAGE ENTRY ON A LINE THAT HAS NO SLASHES (DESIRABLE FOR FIRST RUN WHEN CORRECTION TEXT IS LIKELY TO BE PRESENT).

OPTION BITS FOR SAVEDT .. (IOPT)

- 1) PRINT TRANSLATED INPUT BLKS CLEANED OF EXTRA FEED CHARS.

MEMBER NAME DOCOPERL

- 2) PRINT ERR MSG CONCERNING MTST ENTRY STARTING TOO FAR AFTER BEGINNING OF BLK.
- 4) PRINT ERR MSG CONCERNING LACK OF '1/2' CHAR WITHIN 70 BYTES OF EOB.
- 5) PRINT ERR MSG CONCERNING ILLEGAL(0.9.6) CHARS IN TEXT.
- 6) PRINT MSG CONCERNING START OF EACH NEW SECTION.
- 7) PRINT MAP OF EACH INPUT BLK.
- 8) PRINT EACH PURIFIED TEXT BLK WRITTEN TO 'PURE' FILE.
- 11) PRINT EXPANDED ERR MSG CONCERNING NO MATCHING RIGHT BRACKET.
- 12) PRINT LOG-TYPE ERR MESSAGES ..
(FUNP,BNDX,ICSB,NORX,IXOB,NORB,MTOO).
- 14) ENABLE 'FUNP' ERROR MESSAGES.
- 15) PRINT ERR MSG CONCERNING NONCONSECUTIVE MTST TAPES.
- 17) LIST INDEX ENTRIES PRODUCED.

OPTION BITS FOR STRUCTUR .. (IOP)

- 1) PRINT EACH RAW INPUT SECTION.
- 2) PRINT ONE LINE SECTION HEADER MESSAGE.
- 3) PRINT MAP OF INPUT BLK WITH COORDINATES, ETC FOR LOCATING ERRORS.
- 4) PRINT EACH SECTION OF RAW OUTPUT.
- 5) SCANNER TYPE INPUT, WITH '*' INSTEAD OF PREFIX FOR SECTION DLM.
- 6) PRINT MSG AFTER SUCCESSFULLY PROCESSING LONDON STAGE PAGE ENTRY.
- 7) CREATE UNKEYED SEQ OUTPUT, KEY DATA APPENDED TO FRONT OF BLKS.
- 8) PUT OUT DIRECT ACCESS KEYED BLKS.
- 9) ACCEPT MONTHS IN LADDER REFERENCES IN ANY COMBINATION OF UPPER OR LOWER CASE (USED IN CONJUNCTION WITH ALL UPPER CASE MONSTR CARDS).
- 21) LIST EACH ITEM RETURNED BY GITM.
- 25) SKIP INPUT BLK OPTION.
- 27) SET INDICATOR TO CAUSE PRINTING OF MAP AND RAW OUTPUT FOR SECTIONS ASSOCIATED WITH ERRORS.

OPTION BITS FOR LADDER ..

- 1) PRINT MAP WITH COORDINATES OF EACH INPUT BLK.
- 2) CLEAR LADA FILE AT START OF EACH NEW SEASON. USED IN ORDER TO PROCESS UNLIMITED SIZE FILES ON TAPE.
- 3) SKIP INPUT BLKS OPTION (ALSO IGNORES SEASON CHANGES). FOR DEBUGGING OR SKIPPING AROUND SECTIONS THAT CAUSE PROGRAM TO BOMB ETC.
- 4) PRINT EACH BLK OF RAW INPUT READ DURING LADDER SEARCH.
- 5) ONE LINE HEADER FOR EACH INPUT BLK, GIVING LENGTH, SECTYPE, THEATRE, BLK NUMBER, DATE, AND 1ST PART OF BLK ITSELF.
- 6) SEQUENTIAL FORMAT INPUT FROM FILE STRS.
- 7) SEQUENTIAL FORMAT OUTPUT TO FILE LADS.
- 17) PRINT OLD AND NEW TITLES AFTER LADDER REFERENCE READ-IN.
- 24) PRINT EACH BLK OF RAW OUTPUT.
- 27) SET INDICATOR FOR PRINTING OF EXTRA INFO FOR SECTIONS WITH ERROR MESSAGES.

LADDR HAS A DEBUG CHAIN IN COLS 33-36.

OPTION BITS FOR ITEMGET .. (IOP)

- 1) PRINT MAP OF EACH INPUT SECTION IN LINES OF 100 CHARS EACH WITH COORDINATES FOR EASY LOCATION OF CHARACTERS IN ERROR.
- 2) PRINT EACH ITEM AS IT IS ENCOUNTERED IN INPUT/SETUP. (DEBUG).
- 3) PRINT EACH GRP AS IT IS ENCOUNTERED IN INPUT SETUP. (DEBUG)
- 4) SECTION TITLES IN CAPS.
- 5) PRINT HEADER FOR EACH INPUT SECTION, INCLUDES INPUT SECTION NUMBER, KEY UNTIL IT'S TAKEN OUT, LENGTH OF SECTION, AND FIRST FEW CHARS OF SECTION ITSELF.
- 7) ALLOW FOR PERFORMANCE SELECTION VIA LEVEL 4 INCLUSIVENESS.

MEMBER NAME DOCOPERL

- 8) UNKEYED SEQUENTIAL INPUT FROM LADS FILE INTENDED FOR TAPE INPUT).
- 9) UNKEYED SEQUENTIAL OUTPUT TO ITSM FILE (INTENDED FOR TAPE OUTPUT).
- 10) PRINT CONTROL FLOW TRACE FOR SELECTED TRANSFERS.
- 11) PRINT THE SELECTION CONTROL BLK BEFORE EACH EXIT FROM THE CONTROL CARD INTERPRETER IN NEAT FORMAT WITH HEADINGS.
- 12) CLEAR STATS ARRAY AFTER EACH PRINTING OF IT.
- 19) PRINT EACH OUTPUT RECORD IN FORMAT WHICH USUALLY FITS ON 1 LINE.
- 29-32) ALLOW ERROR MESSAGE FOR ILLEGAL CHARACTER TO BE PRINTED THEATRE, TITLE, ROLE, AND ACTOR RESPECTIVELY.

OPTION BITS FOR FORMAT (IOPt) ..

- 1) LIST IN DATA DIRECTED FORMAT ALL VARIABLES IN PROGRAM AFTER EACH RECORD IS PRINTED (DEBUG).

OPTION BITS FOR LSPRINT .. (IOPt)

- 64) GET DATA TO BE LISTED FROM FILE SYSIN RATHER THAN FILE IN.
- 65) NO PRINTER CARRIAGE CONTROL CHARACTER.
- 66) TYPE NUMBER OF LINES COUNTED ON 1052 AFTER EACH ENDFILE OR END OF RANGE.
- 69) PUT LINE NUMBERS IN F(7) FORMAT STARTING IN COL(83) OF OUTPUT.
- 71) DO NOT COUNT ENTIRELY BLANK LINES (CARRIAGE CTRL CHAR IS CONSIDERED HERE).
- 72) CREATE BACKUP FILE OF INPUT.
- 73) NO TRANSLATION FOR FIRST PRINTING OF LINE.
- 74) NO OVERPRINT ON TOP OF FIRST LINE.
- 75) ENABLE STANDARD PL/1 SYSTEM ACTION FOR ENDPAGE(SYSPRINT) ON-CONDITION, I.E. DON'T IGNORE ENDPAGES.

LSPRINT'S OPTION CARD FOLLOWS THE TRANSLATION TABLES. THERE IS NO OPTION CHAIN, AND LSPRINT CHECKS FOR NO ERRORS.

PLIST OPTIONS

PLIST HAS ONLY 1 OPTION .. LIST LENGTH OF INPUT RECORDS. THIS LENGTH LISTING IS SPECIFIED BY PUTTING ONE OR MORE CHARACTERS IN THE PARM FIELD OF THE EXEC STATEMENT. THERE SHOULD BE NO PARM FIELD IF A PLAIN LISTING IS DESIRED.

ICIFRONT ..

- 1) NO '*' AT LOGICAL BEGINNING OF CORRECTION STATEMENT.
WHEN LOOKING FOR THE '*' THAT DEFINES THE START OF EACH CORRECTION STATEMENT, THE PROGRAM FOUND ANOTHER NONBLANK CHARACTER FIRST. THIS IS A FLUSHING ERROR.

2) ILLEGAL END OF PAGE ENTRY IN LINE N

WHERE N IS THE CURRENT ABSOLUTE LINE NUMBER OF THE DATA TEXT. THIS IS A FAIRLY UNLIKELY ERROR INDICATING THAT SOMETHING IS HORRIBLY WRONG EITHER WITH THE CORRECTION TEXT OR WITH THE PROGRAM. THE PROGRAM STOPS AFTER THIS MESSAGE.

3) NO FCN

AFTER THE BEGINNING OF A STATEMENT NO FUNCTION OR COMMAND CHARACTER CANDIDATES COULD BE FOUND AT ALL. A FATAL ERROR.

4) BAD PAGE SEQUENCE M,N IN LINE X.

PAGE N FOLLOWS PAGE M IN THE DATA TEXT. PAGE N STARTS ON ABSOLUTE LINE X.