

บทที่ 14 : Matching Record

การประมวลผลข้อมูลมีหลายแบบด้วยกัน บางแบบก็อาศัยข้อมูลที่ป้อนเข้ามาเพียง file เดียว บางแบบก็อาศัยข้อมูลที่ป้อนเข้ามามีหลาย file

การประมวลผลข้อมูลโดยอาศัย Input File หลาย File นั้นเราแบ่งออกได้เป็น

1. Merging
2. Matching
3. Selection

1. Merging : คือกระบวนการที่นำหลาย ๆ Sequential Input File มาประมวลผลข้อมูลกันเป็น Sequential Output File อันเดียว ตัวอย่างเช่น การคิดภาษีของพลเมืองแต่ละคนจะมีวิธีการคิดโดยที่บริษัทหรือหน่วยงานแต่ละแห่งจะเก็บประวัติเรื่องราวได้ของคนทำงานในหน่วยงานของตนในแต่ละเดือนโดยเรียงลำดับ พอสิ้นปี ก็จะส่งประวัติรายได้ของทุกคนไปให้รัฐ เพื่อจัดการภาษี รัฐก็จะจัดการนำ Data File ของแต่ละบริษัทแต่ละหน่วยงานซึ่งกระบวนการทำโดยการนำข้อมูลทั้งหมดมารวมกัน แล้วเรียงลำดับตามลำดับของแต่ละหน่วยงาน ซึ่งกระบวนการเช่นนี้จะได้ออก Sequential File อันใหม่

2. Matching : คือกระบวนการที่นำ Input File 2 อันมาเปรียบเทียบกันเพื่อที่จะประมวลผลสำหรับ Record ที่ให้เงื่อนไขเหมือนกัน เช่นรหัสตรงกัน ตัวอย่างเช่น ถ้านำ Input File ที่ประกอบด้วยยอดหนี้เก่าของลูกค้าแต่ละคนของบริษัท มาประมวลผลคู่กับ Input File ซึ่งเป็นข้อมูลในการชำระหนี้ของลูกค้าในเดือนนี้ เราก็จะได้ข้อมูลยอดเงินคงค้างของลูกค้าแต่ละคน ซึ่งเป็น Output File อันเกิดจากการประมวลผลโดยข้อมูล Input File 2 File วิธีการนี้เรียกว่า Matching

3. Selection : คือกระบวนการที่จะเลือกข้อมูลที่ตรงกับวัตถุประสงค์ที่ต้องการมาประมวลผล

ตัวอย่างการใช้ Matching Record

Matching Record เป็นวิธีการที่มีประโยชน์ในการประมวลผลข้อมูลหลายประการที่สำคัญก็คือการนำไปใช้ในการ Merge กับข้อมูลที่เรียงลำดับแล้ว 2 File ให้เข้าเป็นข้อมูลที่เรียงลำดับชุดใหม่ที่อยู่ในแฟ้มข้อมูลเดียวกัน ตัวอย่างเช่นในโรงงานแห่งหนึ่งจะเก็บ Record การทำงานของคนงานแต่ละคนไว้ พอสิ้นสุดวันหนึ่ง ๆ ก็นำ Data File ซึ่งอยู่ในรูปแบบของบัตรนั้นมา Merge กับข้อมูลการทำ

งานของคนงานซึ่งสะสมกันมานานแล้ว จากวันก่อน ๆ หน้าในสลิปดาห์เดียวกัน ซึ่งข้อมูลที่เก็บ สะสมมาก่อนหน้านี้อาจจะอยู่ในเทป พอสิ้นสุดสลิปดาห์ก็จะได้ Data File ซึ่งประกอบด้วยชั่วโมงการทำงานของคนงานทั้งสลิปดาห์ ซึ่ง Data File นี้ก็คือ Transaction File นั่นเอง

จากตัวอย่างนี้เราสามารถจะเขียนคำสั่งใน F-Form และ I-Form ได้ดังนี้คือ



RPG CONTROL CARD AND FILE DESCRIPTION SPECIFICATIONS

Date _____
 Program _____
 PUNCHING INSTRUCTIONS:

Grabic					
Punch					

 Page 1 2
 Program Identification: 75 76 77 78 79 80

Control Card Specifications

Line	Form Type	Core Size to Complete	Object Output Listing Option	Core Size to Execute	Device	MFOM Reading Sequence	MFOM Writing Sequence	Input Shellings	Output Shellings	Control Sequence	MFOM Data	Number of Print Positions	Alternate Copying Sequence	Model 20	Model 26	File Addition/Unreversed	File Condition, UT-UB
0 1	H																

File Description Specifications

Line	Form Type	Filename	File Type	File Designation	Sequence	File Format	Block Length	Record Length	L/R	Mode of Processing	Length of Key Field or of Record Address Field	Record Address Type	Types of File Organization or Additional Area	Overflow Indicator	Key Field Starting Location	Device	Symbolic Device	Label System	Name of Label Exit	Core Index	Extent Exit for DAM	File Addition/Unreversed	Number of Tracks for Cylinder Overflow	Number of Extents	File Condition, UT-UB
0 2	F	FILE1	20	AF	2	WLD										YAES	YS1007								
0 3	F	FILE2	26	AF	2	WLD										RFADLO	YSR008								
0 4	F	NEWTAPE	0	A	2	WLD										YAES	YS1008								
0 5	F	ARZWFILE	0	A	2	WLD										ARWTEAR	YSK107								



RPG INPUT SPECIFICATIONS

Date _____
 Program _____
 PUNCHING INSTRUCTIONS:

Grabic					
Punch					

 Page 1 2
 Program Identification: 75 76 77 78 79 80

Line	Form Type	Filename	Sequence	Number (1-N)	Record Identifying Indicator	Record Identification Codes	Field Location	Field Name	Control Level (1-16)	Measuring Field or Chainable Fields	Field Record Relation	Field Indicators	Starting Sign Position
						1	2	3					
						Position	Position	Position	From	To	Decimal Positions	Plus Minus	Zerf or Blank
0 1	F	FILE1	99	01									
0 2	1								1	9		SSN	
0 3	1								10	18		ADDRS	
0 4	1								13	15		OTHR3	
0 6	2								17	23		DATE	
0 7	1								7	9		SSN	
0 8	1								10	18		ADDRS	
0 9	7								13	15		OTHR3	
1 0	2								17	23		DATE	

จากคำอธิบายลักษณะของคำสั่งได้ดังนี้คือ

คำสั่งใน F-Form แสดงว่ามี 4 File คือ 2 Input File และ 2 Output File โดยที่ Input File อันที่ 1 ชื่อว่า FILE1 ซึ่งเก็บอยู่ในเทป ในที่นี้ก็คือข้อมูลในเรื่องชั่วโมงการทำงานที่เก็บสะสมในวันก่อน ๆ ในสัปดาห์เดียวกัน ส่วน Input File อันที่ 2 ชื่อ FILE2 เก็บอยู่ในรูปบัตร ข้อมูลนี้ก็คือชั่วโมงการทำงานของคนงานที่สิ้นสุดในวันนี้ (เช่น ถ้าวันนี้เป็นวันศุกร์ FILE2 คือชั่วโมงการทำงานของคนงานในวันศุกร์ FILE1 คือชั่วโมงการทำงานสะสมของคนงานตั้งแต่วันจันทร์ อังคาร พุธ พฤหัส นั้นเอง)

ในที่นี้ FILE1 จะทำหน้าที่เป็น Primary File เพราะถือว่าเป็น File ที่สำคัญและยังถือว่าเป็น Master File ดังนั้น จึงใช้อักษร P ในสคมภ์ที่ 18

ส่วน FILE2 จะกลายเป็น Secondary File ไป จึงใส่อักษรตัว S ในสคมภ์ที่ 18

หมายเหตุ ถ้า Input File มีหลาย File ให้ถือ Master File เป็น Primary File และ File อื่น Secondary File ไป ในบรรดา Input File ทั้งหมดจะมี Primary File ได้เพียง 1 อันเท่านั้น ส่วน Secondary File จะมีได้หลายอัน

เนื่องจากการทำงานของปัญหานี้ก็คือการ Merge 2 File เข้าด้วยกัน ดังนั้นทั้ง 2 File จะต้องมีการเรียงลำดับของ Record ให้อยู่ในระบบเดียวกัน คือ ถ้าเรียงจากมากไปหาน้อย ก็ต้องเรียงจากมากไปหาน้อยทั้ง 2 File หรือถ้าเรียงลำดับจากน้อยไปหามากก็ต้องเรียงลำดับจากน้อยไปหามากทั้งสอง File

Output File ที่ได้จากการทำงานขั้นตอนนี้จะมียู่ 2 File คือ NEWTAPE คือข้อมูลที่ได้จากการทำงานในการประมวลผลข้อมูล แล้วจะเก็บอยู่ในเทป ในที่นี้ก็คือข้อมูลที่เรียงลำดับที่ได้จากการ Merge ของ 2 Input File นั้นเอง ส่วน Output File อันที่ 2 คือ PRINTER นั้นจะเป็นการทำงานโดยการพิมพ์ผลรายงานออกมา

เมื่อมี Input File และ Output File ก็จำเป็นจะต้องมีการอธิบายในแบบฟอร์มที่เกี่ยวข้องคือ I-Form และ O-Form ด้วย

คำสั่งใน I-Form คือ

RPG INPUT SPECIFICATIONS

Date: _____
 Program: _____
 Programmer: _____

Punching Instruction	Graphic					
Punch						

Line	Form Type	Filename	Sequence Number (1-10)	Option (0)	Record Identifying Indicator	Record Identification Codes									Field Location		Field Name	Control Level (1-18)	Matching Fields or Chaining Fields	Field Record Relation	Field Indicators			Starting Sign Position															
						1			2			3			From	To					Plus	Minus	Zero or Blank																
						Position	Net (N) CZ/D Character	Character	Position	Net (N) CZ/D Character	Character	Position	Net (N) CZ/D Character	Character																									
01	I	FZLAL	AA	01																																			
02	I														10	19			M2																				
03	I														10	19	SSN																						
04	I														13	15	DATE																						
05	I														17	23	DATE																						
06	I	FZLAL	AA	01																																			
07	I														10	19	SSN																						
08	I														10	19	SSN																						
09	I														13	15	DATE																						
10	I														17	23	DATE																						
11	I																																						
12	I																																						
13	I																																						
14	I																																						

รายละเอียดของแต่ละคำสั่งผู้อ่านคงจะพอเข้าใจได้ เพราะไม่มีอะไรแตกต่างไปจากตัวอย่างที่เคยยกมาให้ดูแล้ว ในส่วนที่แตกต่างที่เพิ่มเติมเข้ามา ก็มีสแตมพ์ที่ 61-62 เป็นส่วนของ Matching-Field ในแต่ละ Input File จะมีที่ใช้ Matching Field อยู่ 2 ส่วนคือ ใน Data Field SSN ใช้ M2 และ DATE ใช้ M1 นั้นหมายความว่าแต่ละ Record ในแต่ละ Input จะมี 2 Field ที่ทำหน้าที่ในการจับคู่ (Matching) กัน

ในตัวอย่างที่ให้นี้เผชิญ ทั้ง 2 Data Field ในต่าง File กันซ้อนตรงกัน แต่จริง ๆ แล้วไม่จำเป็นต้องใช้ชื่อเหมือนกันก็ได้ในการจับคู่ Field การใช้วิธีการจับคู่ Field ก็มีความหมายเหมือนกับการเทียบข้อมูลใน Field ที่จับคู่กันว่ามีข้อมูลเหมือนกันหรือไม่

กฎเกณฑ์ของการจับคู่ Field มีดังนี้คือ

- จำนวน Field สูงสุดที่ใช้ในการจับคู่มีได้แค่ 9 Field เท่านั้นในระบบ DOS (Disk Operating System) แต่ถ้าเป็นระบบ TOS (Tape Operating System) มีได้แค่ 3 Field เท่านั้น
- Field ที่จะนำมาจับคู่กัน จะต้องมีความยาวเท่ากัน และเป็นข้อมูลชนิดเดียวกัน (คือเป็น Numeric หรือ Alpha)
- รหัสที่ใช้ในการจับคู่ของแต่ละ Field ก็คือ M1,M2,...,M9 โดยที่ถือว่า M1 เป็นลำดับที่ต่ำที่สุด และ M9 คือลำดับที่สูงที่สุด

4. ขนาดความยาวของ Field ที่ใช้จับคู่ ไม่ควรเกิน 256 ตัว
5. Field ที่ใช้จับคู่จะปรากฏที่ส่วนใดใน Record ก็ได้ แต่จะต้องอยู่ในตำแหน่งนั้นคงที่ทุก Record
6. จำนวน Field ของแต่ละ Record ในแต่ละ File จะต้องเหมือนกันสำหรับ File ที่นำมาจับคู่กัน
7. ในกรณีที่ Record ของแต่ละ File มีจำนวน Field ที่จะนำมาจับคู่กันมากกว่า 1 Field ขึ้นไป ในการเรียงลำดับของแต่ละ field ในต่าง File กัน ไม่จำเป็นจะต้องเหมือนกัน ตัวอย่างของการจับคู่ในแต่ละ Field ซึ่งไม่เรียงลำดับกัน

IBM International Business Machines Corporation

RPG INPUT SPECIFICATIONS

Date _____
 Program _____
 Programmer _____

Page 1 2
 Program Identification 75 76 77 78 79 80

Line	Form Type	Filename	Sequence Number (LN)	Option (O)	Record Identifying Indicator	Record Identification Codes									Field Location		Field Name	Control Level (L, L1, L2)	Matching Field or Chaining Field	Field Record Relation	Field Indicators			Shelling Sign Position
						1			2			3			From	To					Plus	Minus	Zero or Blank	
						Position	Not (N) CZED Character	Not (N) CZED Character	Position	Not (N) CZED Character	Position	Not (N) CZED Character	Position	Not (N) CZED Character										
01	1	MASTER	AA												1	7	PARTNO	M1						
02	1														8	10	DEPT	M2						
03	1														11	11	INDIV	M3						
04	1														12	14	OPTNAM							
05	1																							
06	1	DETAIL	BB		02	15									1	1	INDIV	M3						
07	1														2	4	DEPT1	M3						
08	1														5	11	MAR71	M1						
09	1														13	14	TSOLD							
10	1																							
11	1																							
12	1																							
13	1																							
14	1																							
15	1																							

ในการทำงานในคอมพิวเตอร์นั้นเมื่อมีการใช้ Matching จากตัวอย่างนี้มี 3 Matching Field คือ M1, M2 และ M3 โดยที่ Matching ที่ 1 มีขนาด Field 1 ตำแหน่ง Matching ที่ 2 มีขนาด field 3 ตำแหน่ง และ Matching ที่ 3 มีขนาด Field 7 ตำแหน่ง ดังนั้นภายใน working area จึงมีการเปรียบเทียบ Field ทั้ง 3 ดังนี้คือ

	M3	M2	M1
Primary (Master)	Division	Department	Post Number
	1	2-4	5-1 1
	M3	M2	M1
Secondary (Detail)	Division	Department	Post Number
	1	2-4	5-11

เงื่อนไขของ Field ที่ใช้ในการจับคู่ เมื่อถูกรับเข้าไปในคอมพิวเตอร์แล้วถ้าหากว่า Record ทั้ง 2 Field มีข้อมูลโดยที่ Field ที่จับคู่ตรงกันแล้ว Matching Indicator เกิดขึ้นโดยที่ Matching Indicator ที่เกิดขึ้นในที่นี้ก็คือ MR ซึ่งจะใช้ MR ในการบังคับเงื่อนไขให้คำสั่งที่ควบคุมในการทำงานจะเกิดขึ้น ดังเช่นใช้ MR ในการคำนวณโปรแกรมตัวอย่างต่อไปนี้

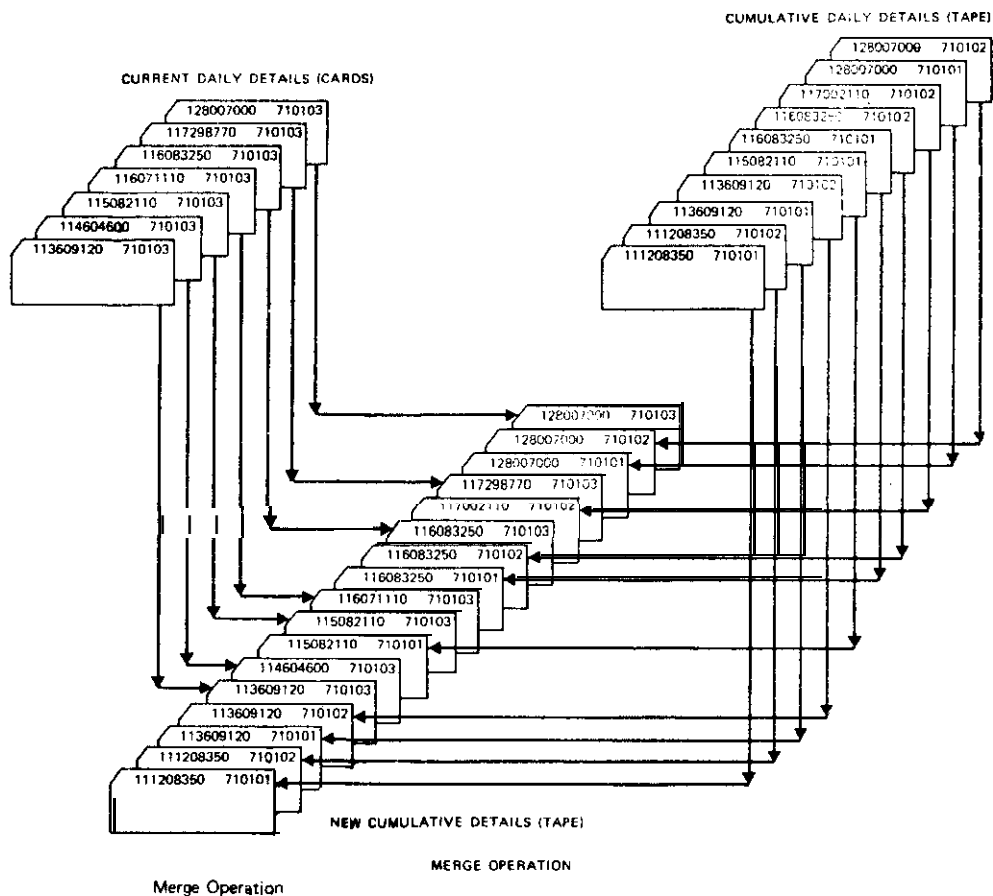
```

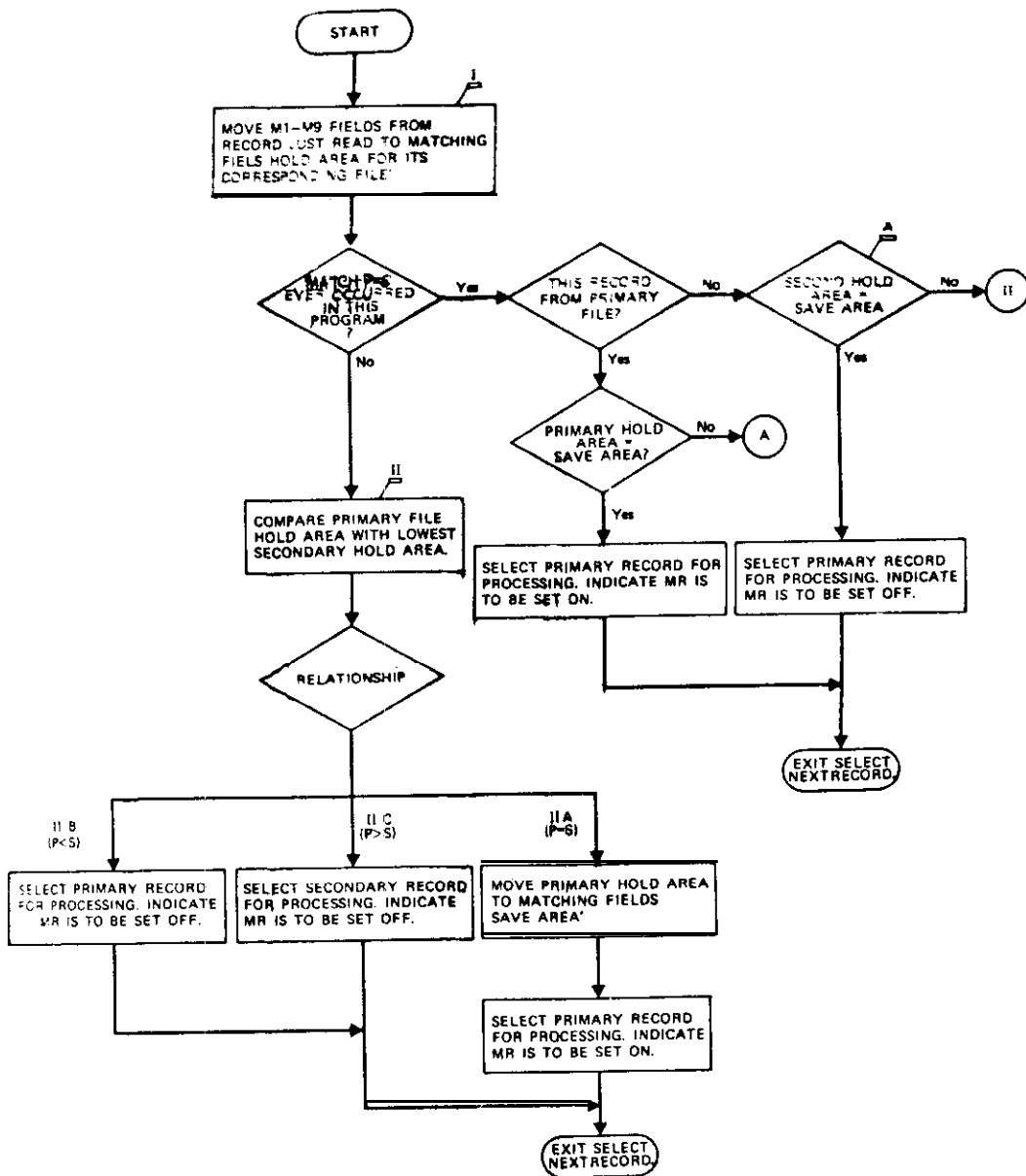
00 C0 C0C H
001 10 020 FFILE1 IP AF 220 22 TAPE SYS007
002 10 030 FFILE2 IS AF 80 80 HEAD40 SYSRDR
003 10 040 FNEWTAPE 0 F 220 22 TAPE SYS008
004 10 050 FPRINTFILE F 132 132 DF PRINTERSYSLST
005 30 010 I* NOTE- FILE1 RECORD NEED NOT BE CHECKED FOR A CODE SINCE ALL
006 30 020 I* RECORDS IN FILE1 ARE VALID.
007 30 030 IFILE1 AA 01
008 30 040 I 1 9 SSN M2
009 30 050 I 10 12 RHRS
010 30 060 I 13 15 OTHRS
011 30 070 I 17 22 DATE M1
012 30 080 IFILE2 BB 02 23 CD
013 30 090 I 1 9 SSN M2
014 30 100 I 10 12 RHRS
015 30 110 I 13 15 OTHRS
016 30 120 I 17 22 DATE M1
017 30 130 I CC 03 23NCD
018 40 010 C 01 CCUNT1 ACC 1 COUNT1 40
019 40 020 C 02 CCUNT2 ACC 1 COUNT2 40
020 40 030 C MR 01 BADCNT ADD 1 BADCNT 40
021 40 031 C* NOTE- WHEN A SECONDARY FILE RECORD MATCHES A PRIMARY FILE RECORD,
022 40 032 C* THIS IS AN ERROR. MR IS TURNED ON, AND MR AND 01 ARE USED
023 40 033 C* TO INCREMENT THE INVALID CARD COUNTER BY 1 FOR EACH SUCH
024 40 034 C* OCCURANCE.
025 40 040 CLR COUNT1 ACC COUNT2 COUNT3 60
026 50 080 CPRINTFILH 201 IP
027 50 090 C CR GF
028 50 093 O 62 *** PAYROLL RECORDS -ERRO*
029 50 094 O 74 *R LISTING ***
030 50 095 O PAGE 2 104
031 50 096 O 100 *PAGE*
032 50 101 C* NOTE- WHEN A MATCH OCCURS BETWEEN A SECONDARY FILE RECORD AND A
033 50 102 C* PRIMARY FILE RECORD, MR IS TURNED ON. MR AND 02 INDICATORS
034 50 103 C* ARE USED TO PRINT OUT THE MATCHING RECORD IN THE SECONDARY
035 50 104 O* FILE.
036 50 105 O D 2 MR 02
037 50 110 O SSN 9
038 50 120 O DATE 20
039 50 125 O 40 *INVALID RECORD*
040 50 130 O T 01 LR
041 50 140 O COUNT3 56 * 0 *
042 50 150 O 50 *TOTAL CARDS PROCESSED=*
043 50 160 O 98 *INVALID CARDS PROCESSED=*
044 50 170 O BADCNT 102 * 0 *
045 52 011 O* NOTE- WHEN NO MATCH OCCURS BETWEEN THE PRIMARY AND SECONDARY
046 52 012 C* FILES, MR IS NOT TURNED ON. NMR INDICATOR IS USED SO THAT
047 52 013 C* THE TWO FILES WILL BE MEGED IN SEQUENCE AND WRITTEN OUT
048 52 014 O* ON TAPE.
049 52 015 C* NOTE- NOT FIRST PAGE IS USED TO ELIMINATE THE INVALID FIRST
050 52 016 C* RECORD THAT WOULD BE WRITTEN ON TAPE DUE TO RPG FIRST
051 52 017 O* RECORD PROCESSING.
052 52 020 CNEWTAPE 0 NMRNIP
053 52 030 O SSN 9
054 52 040 O RHRS 12
055 52 050 O OTHRS 15
056 52 070 O DATE 22

```

โปรแกรมตัวอย่างนี้จะแสดงถึงการใช้ Indicator ในการนับจำนวน Record ซึ่งมีเงื่อนไขของ Field SSN และ DATE ตรงกัน ผู้อ่านคงจะจำได้ว่า ในปัญหาที่เกี่ยวกับการ Merge ข้อมูล ที่เป็น จำนวนชั่วโมงของคณงานนั้น กับจำนวนชั่วโมงสะสม

ตัวอย่างของข้อมูลใน 2 Data File ที่นำมา Merge





Matching Record Processing Logic

การที่จะดึงข้อมูลไปประมวลผลหรือใช้งานนั้น ขึ้นอยู่กับการเก็บรักษาข้อมูลในแฟ้มข้อมูล ซึ่งการเก็บข้อมูลในแต่ละ Record ในแฟ้มข้อมูลนั้นอาจจะเก็บอยู่ในรูปของการเรียงลำดับโดยใช้กฎเกณฑ์หนึ่งกฎเกณฑ์ใดก็ได้ ตัวอย่างเช่น Record ของคนงานแต่ละคนภายในโรงงาน อาจจะจัดเรียงโดยใช้รหัสประจำตัว หรือโดยใช้ชื่อ หรือ Record ของสินค้าแต่ละชนิดเราอาจจะจัดเก็บโดยใช้หมายเลขของสินค้าแต่ละชนิด เป็นหลัก

การจัดเก็บแฟ้มข้อมูลเพื่อไว้ใช้ในหน่วยงานเพื่อให้เป็นแฟ้มข้อมูลที่ี้จะต้องประกอบด้วยองค์ประกอบที่สำคัญ 2 ประการคือ

1. จะต้องเป็นกระบวนการที่ต่อเนื่องกัน (Continuity)
2. ต้องมีการปรับแก้ให้ทันสมัย (Updating)

นั่นหมายความว่ากระบวนการเกี่ยวกับแฟ้มข้อมูลนั้นก็เพื่อที่จะให้ข้อมูลตรงกับความเป็นจริงที่เกิดขึ้นในขณะนั้นและทันกับเวลาที่เปลี่ยนไป ตัวอย่างเช่นเกรดของนักศึกษาจะต้องทันสมัยกับภาคการศึกษา บุคคลในครัวเรือนจะต้องมีรายการตรงกับความเป็นจริงในช่วงเวลานั้น ๆ เช่น การเกิดการตาย และการย้ายที่อยู่จะต้องมีการแจ้งให้ปรับแก้

การที่จะให้แฟ้มข้อมูลมีประสิทธิภาพ ในการนำไปใช้นั้นมีองค์ประกอบที่สำคัญ 2 สิ่งที่ได้กล่าวไว้แล้วนั้นเป็นสิ่งที่พึงระลึกถึงเสมอในการบริหารงานแฟ้มข้อมูล

ความหมายของกระบวนการเกี่ยวเนื่องกับสิ่งต่อไปนี้

- New Record are Inserted
- Old Records are Deleted
- Customer Balances are Updated
- Accounts Payable are Credited or Debited
- Interest Payments must be Made
- Personal Move or Marry
- Equipment Depreciates or Become Obsolete

ความแตกต่างระหว่างแฟ้มข้อมูล (Files) และ Record

แฟ้มข้อมูล (File) คือกลุ่มของ Record ซึ่งมีลักษณะข้อมูลในทิศทางเดียวกัน และจุดมุ่งหมายเดียวกัน ตัวอย่างเช่น แฟ้มข้อมูลเงินเดือนของพนักงาน แฟ้มข้อมูลของลูกค้าของบริษัท แฟ้มประวัติของนักศึกษา แฟ้มข้อมูลบัญชีเงินเดือนของข้าราชการ แฟ้มข้อมูลบัญชีถือฝากของลูกค้าของธนาคาร เป็นต้น

ลักษณะของแฟ้มข้อมูลโดยทั่วไป ไม่ว่าจะเป็นข้อมูลในเรื่องอะไร ก็จะมีการจัดเก็บในลักษณะนี้คือ

Heading : Beginning of File

Detail : Record to be process

Detail

Trailer : End of File

Record คือกลุ่มของข้อมูลซึ่งประกอบด้วยรายการต่าง ๆ (Item) ที่แสดงความสัมพันธ์ร่วมกันในหน่วยใดหน่วยหนึ่ง

ตัวอย่างเช่น ในโรงงานแห่งหนึ่ง จะต้องมีการเก็บประวัติคนงานทุกคน สมมุติว่า ประวัติของคนงานแต่ละคนในโรงงานประกอบด้วย

ชื่อ

รหัส

แผนก

สถานภาพสมรส

เงินเดือน

ดังนั้นประวัติของคนงาน 1 คน ก็คือ 1 Record ซึ่งประกอบด้วยข้อมูล 5 รายการ วิธีเก็บข้อมูลในแฟ้ม จึงขึ้นอยู่กับการวางรูปแบบ และขึ้นอยู่กับตัวกลางที่จะเก็บข้อมูลด้วยว่าจะใช้ตัวกลางใดจึงจะเหมาะสมกับปัจจัยด้านต่างที่เป็นเครื่องกำหนด เช่น ความทนทาน ปริมาณที่จะเก็บ ค่าใช้จ่ายที่จะต้องจัดเก็บ ตลอดจนจนถึงการนำมาใช้งานเมื่อยามต้องการก็จะต้องสะดวกและรวดเร็ว

กรรมวิธีการดำเนินงานเกี่ยวกับเพิ่มข้อมูล

ความสำคัญของการเก็บรักษาเพิ่มข้อมูลเป็นสิ่งที่สำคัญมาก โดยเฉพาะเพิ่มข้อมูลที่มีขนาดใหญ่ ซึ่งจะมีผลต่อการใช้งาน เพิ่มข้อมูลเป็นสิ่งที่สำคัญมากในทุกหน่วยงาน เพราะเพิ่มข้อมูลก็คือแหล่งที่จะแสดงถึงพัฒนาการของหน่วยงานนั้น ๆ และยังแสดงถึงความสัมพันธ์ของข้อมูลภายในหน่วยงานแต่ละแห่งอีกด้วย

โดยทั่วไปแล้วการเก็บเพิ่มข้อมูล ไม่ว่าจะเป็นหน่วยงานเอกชนหรือของรัฐก็ตามจะต้องประกอบด้วยองค์ประกอบที่สำคัญ 2 ประการคือ

1. การเก็บรวบรวม (Storage) มีวิธีการเก็บรวบรวมข้อมูลที่เหมาะสมไม่ว่าจะเป็นการเก็บข้อมูลชนิดชั่วคราว หรือแบบถาวร

2. รวดเร็วทันต่อการนำไปใช้งาน (Accessibility)

หมายความว่าเมื่อไรก็ตามที่ต้องการข้อมูลเพียงบางส่วนในเพิ่มข้อมูลเพื่อนำไปใช้งานก็จะสามารถดึง (Retrive) ไปใช้งานได้ทันที