2. กรณีย์ที่คำตอบว่างเพราะผู้ตอบไม่สามารถตอบได้เพราะไม่มีข้อมูล เช่นคำถามถามว่า พ่อบ้านมีรายได้พิเศษหรือไม่จากการทำงาน ถ้าผู้ตอบเป็นแม่บ้านก็อาจจะตอบไม่ได้เพราะไม่ทราบก็ได้ หรือตัวอย่างคำถามประเภทที่ผู้ตอบไม่มีความรู้เรื่องนั้น เช่น ถามว่าท่านเห็นด้วยหรือไม่กับการเลือก ส.ส. แบบแบ่งเขต คำถามเช่นนี้ถ้าผู้ตอบ ไม่มีความรู้ในเรื่องนี้ คำตอบในข้อดังกล่าวจะว่างไป

3. กรณีย์ที่คำตอบว่างเพราะผู้ตอบไม่ต้องการตอบอันอาจเนื่องมาจากความละอายใจที่จะตอบ ตัวอย่างเช่น คำถามเรื่องเกี่ยวกับการคุมกำเนิดถ้าผู้ตอบเป็นสตรีก็มักจะละอายใจและพาลไม่ตอบคำถามเลย ดังนั้นข้อมูลในคำถามดังกล่าวจะว่างไป

ดังนั้นเมื่อผู้ลงระหัสเห็นข้อมูลที่ว่าง ๆ ก็จะต้องมีการแยกแยะระหัสที่จะลงไปเพื่อที่จะได้มีการทราบถึงความแตกต่างของข้อมูลที่ว่างดังกล่าวทั้งนี้เพราะบางครั้ง ถ้าหากพบข้อมูลในรายการใดรายการหนึ่งมีการว่างมากผิดปกติก็สามารถแยกวิเคราะห์ข้อที่ว่างเนื่องมาจากสาเหตุใด จะต้องมีการแก้ไขหรือไม่เป็นต้น โดยปกติมักจะมีการให้ระหัสเป็นสากลกัน ดังนี้คือ จะให้ระหัสเลข

9. แทนการว่าง เพราะไม่ทราบ (don't know)

8. แทนการว่าง เพราะไม่ตอบ (No answer)

7. แทนการว่าง เพราะไม่ต้องตอบ (does not apply)

กระบวนการลงระหัสดังกล่าวเป็นขั้นตอนที่ใช้คนมาก ในการลงระหัส ซึ่งจะต้องมีการอบรมทำความเข้าใจในวิธีการให้เข้าใจเสียก่อน ซึ่งถ้าแม้จะอบรมแล้วแต่ก็จะต้องจัดทำคู่มือในการลงระหัสให้ใช้ด้วย เพื่อจะได้ช่วยในการลงระหัสให้สะดวกยิ่งขึ้นและจะได้ไม่สับสนเวลาลงระหัสจริง ๆ คู่มือลงระหัสที่เรียกว่า Coding Book
ความหมายของ Coding Book ก็คือ 'The code-book is a dictionary where questions or variables are translated into columns and answers or values are translated into punches'
ตัวอย่างของ Coding Book ในงานวิจัยชั้นหนึ่ง

คู่มือลงระหัส

การติดตามผลบัณฑิตคณะวิทยาศาสตร์ มหาวิทยาลัยรามคำแหง

ข้อที่	สดมภ์ที่
1. ให้ลงระหัสของแบบสอบถาม	1-3
2. เพศ ——— " ———	4
ตอบ หญิง เจาะ 1	
ชาย เจาะ 2	

จากตัวอย่าง Coding Book ที่แสดงให้ดูจะเห็นว่าใน Coding Book จะต้องบอกถึงสดมภ์ที่จะต้องเจาะด้วย ดังนั้นผู้เขียนคู่มือลงระหัสจะต้องวางรูปแบบของข้อมูลลงในบัตรชนิด 80 สดมภ์ก็จะต้องมีการ design card format ด้วยการบันทึกข้อมูลลงตัวกลางชนิดอื่นก็เช่นกันมีวิธีการบันทึกเช่นเดียวกัน เพื่อให้ผู้ที่ยังไม่เคยทราบถึงวิธีการบันทึกข้อมูลลงตัวกลางเลย จะขอยกตัวอย่างโดยการใช้บัตรชนิด 80 สดมภ์ เป็นตัวอย่างเพื่อแสดงถึงวิธีการวางแบบข้อมูล (design card format) และวิธีการบันทึกระหัสลงในบัตร

ตัวกลางที่ใช้เก็บข้อมูล (Media) เนื่องจากว่าเครื่องคอมพิวเตอร์ไม่สามารถอ่านตัวหนังสือที่อยู่ในเอกสารได้ จึงต้องมีการถ่ายทอดข้อมูลจากเอกสารเบื้องต้นให้อยู่ในรูปแบบที่เครื่องสามารถอ่านได้ (Machine Readable Form) ซึ่งตัวกลางที่บรรจุข้อมูลที่มีคุณสมบัติดังกล่าวแยกได้เป็นหลายชนิดเช่น

1. บัตรเจาะรู (Punched Card)
2. เทปแม่เหล็ก (Magnetic Tape)
3. จานแม่เหล็ก (Magnetic Disk)
4. ดรัม (Magnetic Drum)
5. เทปกระดาษ (Paper Tape)
6. ดาต้าเซล (Data Cell)
7. ดิสเคท (Diskette)

ซึ่งในบรรดาตัวกลางเหล่านี้ จะขอยกมากล่าวเพียงบางตัวที่สำคัญ ๆ และใช้กันมาก

บัตรเจาะรู (Punched Card)

ผู้ที่คิดค้นบัตรเจาะรูขึ้นมาใช้เป็นครั้งแรกคือ Joseph Jacquard ชาวฝรั่งเศส จุดประสงค์ในการคิดค้นนี้ก็คือ การนำเอาบัตรดังกล่าวมาใช้ควบคุมลายในการทอผ้า เพื่อให้ได้ลวดลายของผ้าตามที่ต้องการ ต่อมา Dr. Herman Hollerith ได้นำความคิดนี้มาใช้กับ

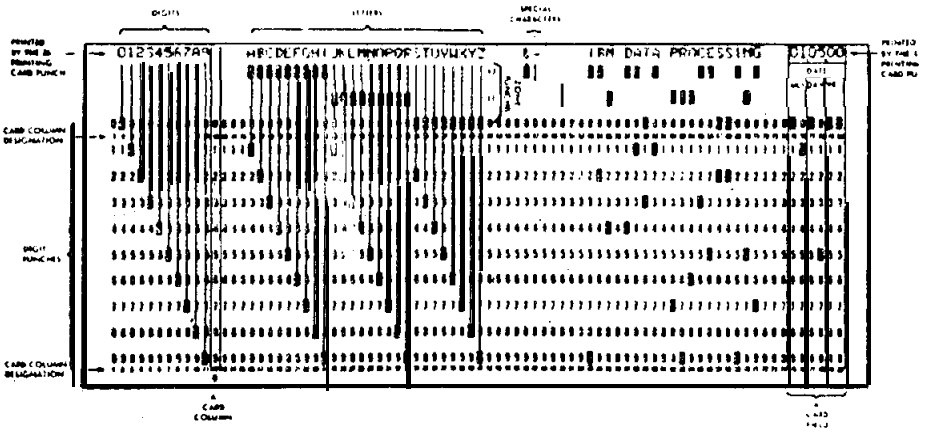
บัตรเจาะรูเพื่อใช้ในการประมวลผลข้อมูลในการสำมะโนประชากรของสหรัฐอเมริกาในปี ค.ศ. 1890 บัตรเจาะรูนี้เรียกว่า Hollerith Card (เรียกกันภายหลังว่าบัตร IBM) นอกจากจะคิดค้นใช้บัตรเจาะรูแล้วก็ยังมีการคิดค้นเครื่องจักรกลที่ใช้ในการจำแนกบัตรด้วย

ลักษณะของบัตรเจาะรู

บัตรเจาะรูซึ่ง Hollerith ออกแบบในปี ค.ศ. 1887 มีลักษณะเป็นบัตรแข็งเป็นรูปสี่เหลี่ยมผืนผ้าขนาดยาว $7\frac{3}{4}$ นิ้ว ความกว้าง $3\frac{1}{4}$ นิ้ว ถึงแม้ปัจจุบันนี้ ลักษณะของบัตรที่ใช้กันอยู่ก็ยังคงเดิม บัตรดังกล่าวแบ่งเป็น 80 คอลัมน์ แต่ในปัจจุบันนี้มีบัตรชนิดใหม่มีลักษณะแตกต่างไปอีกหลายแบบ เช่นเป็นบัตรที่มี 96 คอลัมน์และมีลักษณะเป็นรูปสี่เหลี่ยมจัตุรัส สำหรับใช้กับเครื่อง System 3 และยังมีบัตรชนิด 90 คอลัมน์ แต่ไม่ค่อยนิยมใช้กันมากเหมือนบัตรชนิด 80 คอลัมน์และชนิด 96 คอลัมน์

ลักษณะของบัตร 80 สดมภ์ (คอลัมน์) และการบันทึกข้อมูล

บัตร 80 สดมภ์เป็นที่นิยมใช้กันมากและใช้กันทั่วไปในปัจจุบันนี้ ลักษณะของบัตรชนิดนี้เป็นบัตรแข็งรูปสี่เหลี่ยมผืนผ้าแบ่งออกตามความยาวได้ 80 สดมภ์ โดยที่แต่ละสดมภ์จะมีเลขกำกับไว้ตั้งแต่เลข 1-80 ในแต่ละสดมภ์จะแบ่งออกเป็น 12 แถว และใน 12 แถวนี้จะแบ่งออกเป็น 2 ส่วน ส่วนที่เห็นเป็นช่องว่าง (ไม่มีตัวเลขกำกับ) อยู่เหนือแถวที่เป็นเลข 0 จะเป็นที่ตั้งของแถวที่ 11 (X Punch) และแถวที่ 12 เรารวมเรียกแถวที่ 0 แถวที่ 11 แถวที่ 12 นี้ว่า Zone Punching Area ถัดจาก 0 ลงมาข้างล่างประกอบด้วยแถว 1, 2, .:., 9 ซึ่งในแต่ละแถวดังกล่าวจะมีเลขกำกับอยู่ด้วย รวมเรียกแถว 0, 1, 2 ..., 9 ว่า Digit Punching Area จะสังเกตเห็นว่าบัตรแต่ละใบจะถูกตัดริมที่มุมซ้ายออก ซึ่งเรียกส่วนนี้ว่า Upper Right Corner Cut จุดประสงค์ที่ตัดมุมบัตรออกก็เพื่อความสะดวกในการเรียงบัตรไม่ให้กลับตัวกัน (ดูรูปประกอบ)



ในแต่ละสดมภ์ของบัตรแต่ละใบจะสามารถเจาะรูสี่เหลี่ยมเล็ก ๆ ได้ถึง 12 รูด้วยกัน (เพราะว่ามีอยู่ 12 แถว) ลักษณะของการเก็บข้อมูลลงในบัตรดังกล่าวโดยอาศัยตำแหน่งของรูที่เจาะในแต่ละแถวประกอบกันเข้า

ข้อมูลที่ใช้โดยทั่วไปแยกออกเป็น 3 ประเภท

1. ข้อมูลที่เป็นตัวเลข (Numeric Data) ลักษณะการสร้างข้อมูลชนิดนี้ก็ทำโดยการเจาะตัวเลขลงในสดมภ์ที่ต้องการ ตัวเลขก็อาศัยการเจาะรูเท่านั้น เช่นต้องการเจาะเลขในสดมภ์ที่ 1

ก็เจาะรูที่แถวที่ 1 ในสดมภ์ที่ 1 สำหรับตัวเลขอื่นก็เจาะเช่นเดียวกัน ดังนั้นในสดมภ์ที่ 1 ก็จะสร้างตัวเลขได้เพียง 1 ตัวเท่านั้น

2. ข้อมูลที่เป็นตัวอักษร (Alphabetic Data) จะต้องอาศัยการประกอบกันของตำแหน่งต่าง ๆ ของรูที่เจาะโดยมีหลักเกณฑ์ดังนี้

ตัวอักษร A-Z จะแบ่งออกเป็น 3 กลุ่ม

กลุ่มที่ 1. A-I จะใช้การประกอบกันของเจาะรูที่ตำแหน่งแถวที่ 12 (Zone Punch) กับตำแหน่งของตัวเลขแถวที่ 1-4 (Digit Area)

เช่น A เจาะรูที่แถวที่ 12 กับแถวที่ 1

B เจาะรูที่แถวที่ 12 กับแถวที่ 2

I เจาะรูที่แถวที่ 12 กับแถวที่ 9

กลุ่มที่ 2. J-R จะใช้การประกอบกันของเจาะรูที่ตำแหน่งที่ 11 (Zone Punch) กับตำแหน่งของตัวเลขในแถวที่ 1-9 (Digit Area)

เช่น J เจาะรูที่แถวที่ 11 กับแถวที่ 1

K เจาะรูที่แถวที่ 11 กับแถวที่ 2

.

R เจาะรูที่แถวที่ 11 กับแถวที่ 2

กลุ่มที่ 3. S-Z จะใช้การประกอบกันของการเจาะรูที่ตำแหน่งแถวที่ 0 (Zone Punch) กับตำแหน่งของตัวเลขในแถวที่ 1-9

เช่น S เจาะรูที่แถวที่ 0 กับแถวที่ 2

T เจาะรูที่แถวที่ 0 กับแถวที่ 3

Z เจาะรูที่แถวที่ 0 กับแถวที่ 9

ข้อสังเกต S จะไม่ใช่การประกอบกันของแถวที่ 0 กับแถวที่ 1 แต่ใช้แถวที่ 2 แทน

3. ข้อมูลที่เป็นสัญญลักษณ์พิเศษ (Special Character)

เช่นพวกเครื่องหมายทางคณิตศาสตร์ต่าง ๆ และสัญญลักษณ์พิเศษอื่น ๆ ให้ดูรายละเอียดในการเจาะจากตาราง

CHARACTER	PUNCH POSITION	CHARACTER	STANDARD	EXTENDED
I	1	⟨		12, 4, 8
2	2	⟩		0, 6, 8
3	3	+	12	12, 6, 8
4	4	-		0, 5, 8
5	5	)	12.4, 8	11, 5, 8
6	6	(	0, 4, 8	12, 5, 8
7	7	¢		12, 2, 8
8	8			2, 8
9	9			11.6, 8
0	0	Π		11, 7, 8
	11		4, 8	5, 8
&	12	?		0, 7, 8
	12, 3, 8	"		7, 8
$	11, 3, 8	=	3, 8	6, 8
	0, 3, 8	!		11, 2, 8
#	3, 8	⌑		12, 7, 8
•	11, 4, 8		12, 4, 8	
%	0,4, 8			
@	4, 8			
/	0, 1			

ในปัจจุบันนี้ประเทศไทยมีเครื่องเจาะบัตร ซึ่งสามารถเจาะตัวอักษรเป็นภาษาไทย โดยอาศัยการประกอบกันของการเจาะรู ณ. ตำแหน่งต่าง ๆ โดยที่การเจาะ 3 รูในสดมภ์เดียวกันจะได้พยัญชนะหรือสระในภาษาไทยได้ 1 ตัว

รายการข้อมูล (Data Item)

รายการข้อมูลก็คือกลุ่มของ Character ตั้งแต่ 1 ตัวขึ้นไปที่มีความเกี่ยวพันกันในความหมายใดความหมายหนึ่ง ตัวอย่างของรายการข้อมูล (Data Item) เช่น ชื่อคนงาน ระหัสของคนงาน จำนวนเงิน เป็นต้น

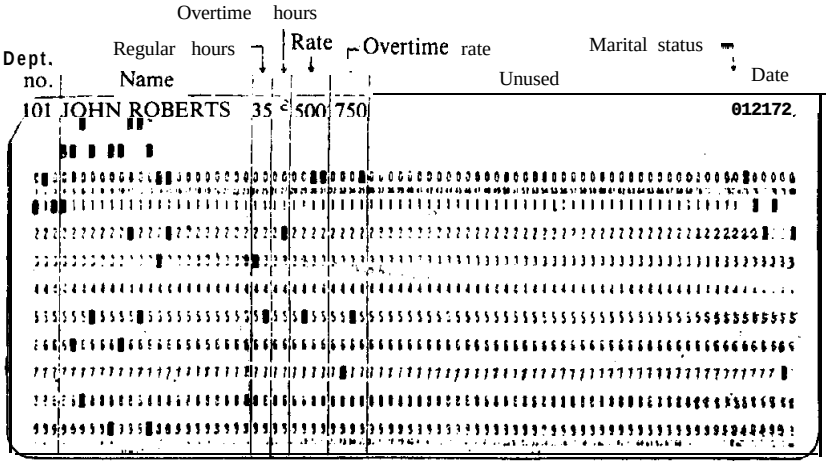
Card Field

Field คือพื้นที่ส่วนหนึ่งในบัตรซึ่งคิดจากจำนวนสดมภ์ที่ประกอบกันอยู่ Field หนึ่ง ๆ อาจจะประกอบด้วย อย่างน้อยที่สุด 1 สดมภ์หรืออย่างมากที่สุด 80 สดมภ์ (เมื่อใช้บัตรชนิด 80 สดมภ์)

จุดประสงค์ของการสร้าง Field ก็เพื่อที่จะนำไปใช้ในการเก็บข้อมูลหนึ่ง ๆ (Data Item)

ให้พิจารณาการสร้าง Card Field เพื่อเก็บข้อมูลแต่ละรายการจากเอกสารเบื้องต้น (Source Document) เพื่อประกอบความเข้าใจยิ่งขึ้น

Name JOHN ROBERTS
Dept. 101
Regular hours 35
Overtime 2
Rate 5.00
Overtime rate 7.50
Week ending 1/21/72
D. Richards
Supervisor



Payroll card

Field name	Card column (s)
1. Department number	l-3
2. Name	4-23
3. Regular hours	24-25

4. Overtime hours	26-27
5. Rate	28-31
6. Overtime rate	32-35
Unused	36-73
7. Marital status	74
8. Pay period date	75 -80

ข้อสังเกต การเจาะข้อมูลชนิดที่เป็นตัวอักษรล้วน เช่น ชื่อคนโดยปกติเราจะเจาะชิดซ้ายของ Field นั้น (Left Justified) ดูจากตัวอย่างใน Field ที่ 2

ส่วนการเจาะข้อมูลชนิดที่เป็นตัวเลขล้วน ๆ เราจะเจาะชิดขวาของ Field นั้น ๆ (Right Justified) ดูตัวอย่างใน Field ที่ 3, 4, 5 เป็นต้น

ในบางกรณีข้อมูลในรายการอาจจะเป็นทั้งตัวอักษรและตัวเลขรวมกันอยู่ เช่น รหัสของคนงานเป็น A1213 เรียกข้อมูลชนิดนี้ว่า Alphameric

วิธีการสร้าง Card Field เพื่อใช้กำหนดจำนวนสดมภ์ที่จะเจาะข้อมูลเป็นส่วนหนึ่งที่เรียกว่า Card Design ซึ่ง Card Design นับว่าเป็นขั้นตอนที่สำคัญในการเก็บข้อมูลลงในบัตร

Card Design ที่ดีนั้นจะต้องประกอบด้วย

1. Field แต่ละรายการที่สร้างจะต้องเหมาะสมกับข้อมูลแต่ละรายการ นั่นคือข้อมูลในแต่ละรายการจะต้องบรรจุลงใน Field นั้นได้พอดี คือจำนวนสดมภ์จะต้องไม่มากเกินไปหรือน้อยเกินไปสำหรับข้อมูลแต่ละรายการ

ตัวอย่างของ Card Design ที่ไม่ดี

ก. ให้จำนวนสดมภ์ในแต่ละรายการน้อยเกินไปกว่าจำนวนข้อมูลในรายการนั้น ๆ จะบันทึกลงหมดได้ เช่น กำหนดให้ Field ในเรื่องรายได้ของคนในกรุงเทพฯ เท่ากับ 6 สดมภ์แต่ปรากฏว่ามีตัวอย่างรายหนึ่งในกรุงเทพฯ เกิดมีรายได้เป็น 7 หลัก นั้นก็หมายความว่าข้อมูลในเรื่องรายได้ของบุคคลผู้นี้ไม่สามารถบันทึกลงใน Field ที่กำหนดไว้ได้ ทำให้เกิดเสียหายกับข้อมูลที่บันทึกมาตั้งแต่แรก และยังต้องทำให้มีการวาง Card Design กันใหม่

ข. การกำหนดให้จำนวนสดมภ์ใน Field ใด ๆ Field หนึ่งมากกว่าจำนวนข้อมูลในแต่ละรายการตัวอย่างเช่นในเรื่องข้อมูลเรื่องเพศ ซึ่งบ่งไว้ว่าถ้าเป็นเพศชายจะเจาะเลข

1 และถ้าเป็นเพศหญิงจะเจาะเลข 0 ใน Field นั้น จะเห็นได้ว่าจำนวนสดมภ์ใน Field นั้นใช้เพียง 1 สดมภ์ก็เพียงพอแต่ไปกำหนดถึง 3 สดมภ์ ซึ่งจะเสียประโยชน์ไปถึง 2 สดมภ์

การที่เหลือสดมภ์ว่างใน Field นั้นถ้าเป็นกรณีของการประมวลผลข้อมูลที่มีข้อมูลไม่มากนักก็ไม่มีผลอย่างไร แต่ถ้ามีข้อมูลเป็นจำนวนมากแล้ว การเสียแต่ละสดมภ์ไปโดยไม่ใช้ประโยชน์ย่อมหมายถึงการเสียค่าใช้จ่ายไปจำนวนหนึ่งโดยเปล่าประโยชน์ ตัวอย่างที่เห็นได้ชัดก็คือในกรณีของการประมวลผลในงานสำมะโนประชากรของประเทศ ที่มีข้อมูลที่จะประมวลผลเป็นจำนวนมาก ดังนั้นการ Design Card Format จึงถือว่าเป็นเรื่องที่สำคัญมากเช่นกัน

2. Field แต่ละ Field ควรจะเรียงลำดับเป็นไปเช่นเดียวกันกับข้อมูลที่ถ่ายทอดจากเอกสารเบื้องต้น ทั้งนี้เพื่อความสะดวกและง่ายแก่การตรวจสอบความถูกต้องภายหลัง

3. ผู้วางแบบ Card Format ควรจะเขียนรายการสำหรับผู้เจาะบัตร ในรายการดังกล่าวก็จะประกอบด้วย

3.1 ชื่อของ Field (Field Name) ซึ่งจะใช้รายการของข้อมูลในเอกสารเบื้องต้นเป็น

3.2 Card Column คือตำแหน่งของสดมภ์ที่เริ่มต้นของ Card Field ในแต่ละรายการเช่น

Field Name	Card Column
1. Department Number	1-3
2. Name	4-23
3. Regular Hours	24-25

3.3 คำตอบที่เป็นไปได้ในแต่ละรายการของข้อมูล

(All Possible Answers in Each Data Item)

การบอกคำตอบที่เป็นไปได้ก็เพื่อจุดประสงค์ในการตรวจสอบความผิดพลาดของข้อมูล และเพื่อความสะดวกของการเจาะข้อมูล ตัวอย่างเช่น รายการของข้อมูลที่เกี่ยวกับสถานะภาพสมรส คำตอบ (ข้อมูล) ที่จะเป็นไปได้คือ

Field Name	Card Column	Possible Answer
Marital Status	74	1 - Single
		2 - Married

		3 - Withdraw
		4 - Widow
Age	75-76	Actual Age

ลักษณะของ Character ที่เจาะในบัตร 80 สดมภ์นี้เราเรียกว่า Holletirh Code

ลักษณะของบัตร 96 สดมภ์และการบันทึกข้อมูล

บัตร 96 สดมภ์ใช้กับคอมพิวเตอร์ IBM System 3 บัตรแต่ละใบจะแบ่งทางแถวแนวนอนเป็น 2 ส่วน (Zone)

ส่วนที่ 1 เรียกว่า Print Area

ส่วนที่ 2 เรียกว่า Punch Area

ในส่วนที่ 1 คือ Print Area จะแบ่งย่อยออกเป็น 3 Print Line แต่ละ Print Line จะบรรจุได้ 32 Characters ดังนั้น 3 Print Line จะบรรจุได้ 96 Characters

ในส่วนที่ 2 คือ Punch Area จะแบ่งย่อยออกเป็น 3 Punch Area โดยที่แต่ละ Punch Area จะเจาะรูเพื่อสร้างเป็น Character ต่าง ๆ ที่ต้องการได้ถึง 32 Character โดยที่ผลลัพธ์ (ข้อมูล) ที่ได้จากการเจาะ Character ใน Punch Area ส่วนที่ 1 จะไปปรากฏในส่วน Print Area ที่ 1 (ดูรูปประกอบสำหรับรายละเอียดย่อยอื่น ๆ ในตัวบัตร)

เมื่อเปรียบเทียบบัตรชนิด 80 สดมภ์กับบัตรชนิด 96 สดมภ์จะเห็นได้ว่าบัตรชนิด 80 สดมภ์ใช้ 12 แถวในแต่ละสดมภ์เพื่อสร้าง 1 Character แต่ในบัตร 96 สดมภ์ใช้เพียง 6 แถวในแต่ละสดมภ์เพื่อสร้าง 1 Character แต่ละแถวเพื่อที่ประกอบกันขึ้นเป็น 1 สดมภ์ เราใช้สัญญลักษณ์แทนดังนี้คือ

B

A

8

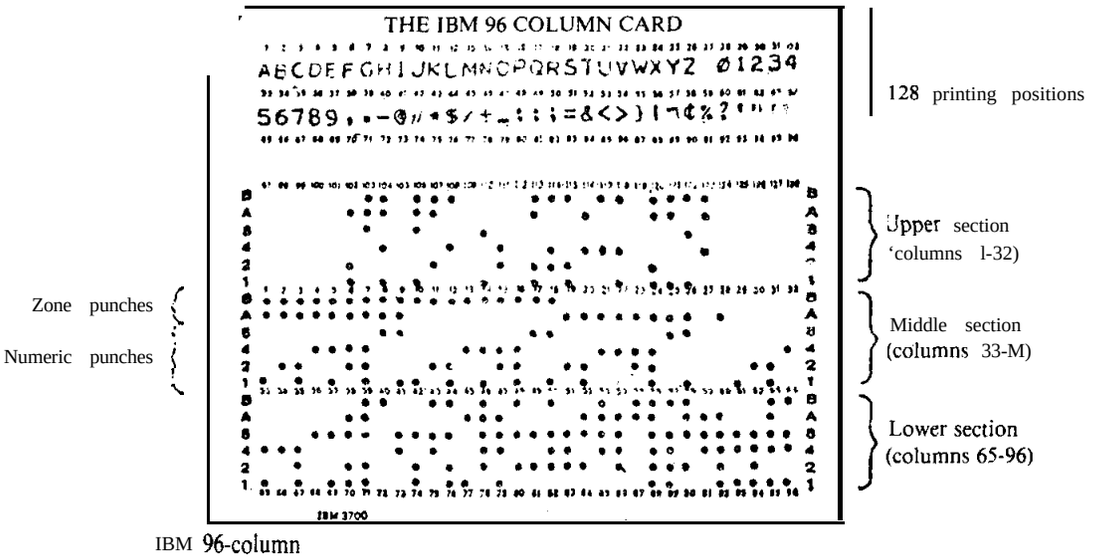
4

2

1

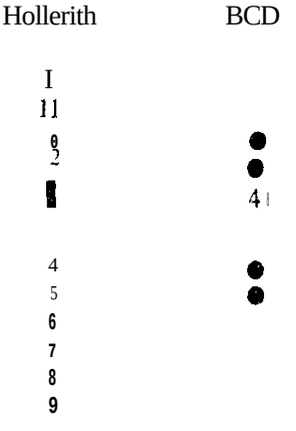
เรียกระบบนี้ว่า Binary-Coded-Decimal (BCD)

ตำแหน่ง B, A ทำหน้าที่คล้ายกับ Zone 12, 11 และ 0 ในระบบ Hollerith Code ส่วนตำแหน่ง 8, 4, 2,1 จะทำหน้าที่คล้ายกับเลข 1, 2, 3, 4, 5, 6, 7, 8, 9 ในระบบ Hollerith Code ดูรูปประกอบจะเข้าใจยิ่งขึ้น



IBM 96-column

รูปต่อไปนี้จะเปรียบเทียบความแตกต่างระหว่างการแทน Character C ด้วยระบบ Hollerith Code และ BCD Code



เปรียบเทียบข้อดีและข้อเสียของการใช้บัตรเป็นตัวกลางเก็บข้อมูล

ข้อดีของการใช้บัตร	ข้อเสียของการใช้บัตร
1. แก้ไขข้อมูลที่ผิดพลาดได้ง่าย	1. ไม่สะดวกในการขนย้ายข้อมูลเพราะมีน้ำหนักมาก
2. ตรวจสอบข้อมูลได้ง่าย (ใช้สายตาได้)	2. เก็บรักษายากยิ่งในกรณีที่อากาศชื้น บัตรอาจจะบวมทำให้ใช้ไม่ได้
	3. เปลืองเนื้อที่ในการเก็บรักษา
	4. ไม่ทนทาน เปื่อยยุ่ยง่าย
	5. ใช้เก็บข้อมูลได้ครั้งเดียวไม่สามารถนำไปเก็บข้อมูลชุดใหม่ได้ (ลบข้อมูลในบัตรชุดเดิมทิ้งไม่ได้)

บัตรที่ใช้เก็บข้อมูลเมื่อเวลาจะส่งเข้าเครื่องคอมพิวเตอร์ เพื่อประมวลผลข้อมูลจะส่งเข้าทาง Input-Unit ซึ่งเรียกว่า Card Reader ซึ่งจะทำหน้าที่ในการอ่านข้อมูลจากบัตรเข้าสู่เครื่องคอมพิวเตอร์ เพื่อนำไปประมวลผล

ข้อสังเกต: การเจาะข้อมูลลงในบัตรไม่ว่าจะเป็นบัตรชนิดใดก็ตาม มีหลักเกณฑ์ดังนี้คือ

1. ถ้าเป็นข้อมูลที่เป็นตัวเลข (Numeric Data) ให้บันทึกข้อมูลลงใน Field ที่ต้องการโดยยึดหลักว่า ให้หลักหน่วยอยู่ที่สดมภ์ทางขวาสุดของ Field นั้น วิธีนี้เรียกว่าเจาะแบบ Right Justified การบันทึกข้อมูลถ้าหากไม่เต็ม Field จะเจาะลงไปทางสดมภ์ทางซ้ายมือที่เหลือ (หรือจะทิ้งว่าง (blank) ไว้ก็ได้)

หมายเหตุ สดมภ์ทางซ้ายสุดของ Field เราเรียกว่า High-Order-Column ส่วนสดมภ์ทางขวาสุดของ Field เรียกว่า Low-Order-Column

1. การบันทึกข้อมูลที่เป็นตัวอักษร (Alphabetic Data) ให้บันทึกข้อมูลแบบชิดไปทางซ้ายของ Field ซึ่งการบันทึกเช่นนี้เรียกว่า Left Justified ส่วนที่ไม่ได้บันทึกของ Alphabetic อยู่ใน Low Order

หมายเหตุ ภายใน Alphabetic Field เช่น Name Field อาจจะมีช่องว่างก็ได้ เช่น WICHAI SUDEE ในกรณีเช่นนี้ช่องว่าง (หลัง WICHAI) ก็ถือว่าเป็นส่วนหนึ่งของข้อมูลที่ต้องการบันทึกได้

เทปแม่เหล็ก (Magnetic Tape)

เทปแม่เหล็กเป็นตัวกลางที่เก็บข้อมูลชนิดหนึ่งแผ่นยาวความกว้างประมาณ $\frac{1}{4}$ นิ้ว ความยาวมีหลายขนาดบรรจุเป็นม้วน (Reel) วัดความจุของเทปแม่เหล็กด้วยความหนาแน่นของจำนวน Character ที่บรรจุลงในเทปแม่เหล็ก 1 นิ้ว โดยใช้หน่วยเรียกว่า BPI (Byte per Inch) โดยปกติแล้วเทปแม่เหล็กทั่ว ๆ ไปจะมีความหนาแน่น 200-1600 Character per Inch เทปแม่เหล็กเป็นได้ทั้ง Input และ Output Media ดังนั้น Magnetic Tape Unit จึงเป็นได้ทั้ง Input และ Output Unit สำหรับ Computer System คือ เป็นได้ทั้งเครื่องอ่านและบันทึก

ข้อมูลที่จะบันทึกลงในเทปแม่เหล็กจะบันทึกอยู่ในรูปของ Magnetized Spot ซึ่งเรียกว่า bit เทปแม่เหล็กขนาดมาตรฐานคือ $10\frac{1}{2}$ inch reel (ความยาวประมาณ 2400 ฟุต) น้ำหนักประมาณ 4 ปอนด์ จุได้ 20 Million Characters ซึ่งเทียบได้เท่ากับบัตรชนิด 80 สดมภ์ (ต้องบันทึกเต็มใบ) ประมาณ 250,000 ใบ เทปแม่เหล็กที่ยาว 2,400 ฟุต และมีขนาดความหนาแน่น (density) 800 Character ต่อนิ้ว แล้วเทปม้วนนี้จะสามารถจุได้ถึง 2400 x 12 x 800 = 23,040,000 Characters (23 ล้าน Characters เทียบได้เท่ากับบัตรชนิด 80 สดมภ์คือ 2,304,000 ÷ 80 = 288,000 ใบหรือประมาณ 144 กล่อง)

ความเร็วในการอ่านข้อมูลจากเทปไปยัง Main Storage ซึ่งเรียกว่า Transfer Rate สูงมากเมื่อเทียบกับการใช้บัตรเป็นตัวกลาง กล่าวคือถ้า Transfer Rate ของการใช้บัตรเป็น 2,667 Characters per seconds แล้ว เทปแม่เหล็กจะมี Transfer Rate ในอัตรา 350,000 Characters per second โดยปกติแล้ว Transfer Rates ของเทปแม่เหล็กขึ้นกับปัจจัยดังต่อไปนี้คือ

1. ความหนาแน่นของเทป (Magnetized Marks)
2. ความเร็วในการหมุนของเทป ปกติเทปจะหมุนความเร็วประมาณ 100 นิ้วต่อวินาที

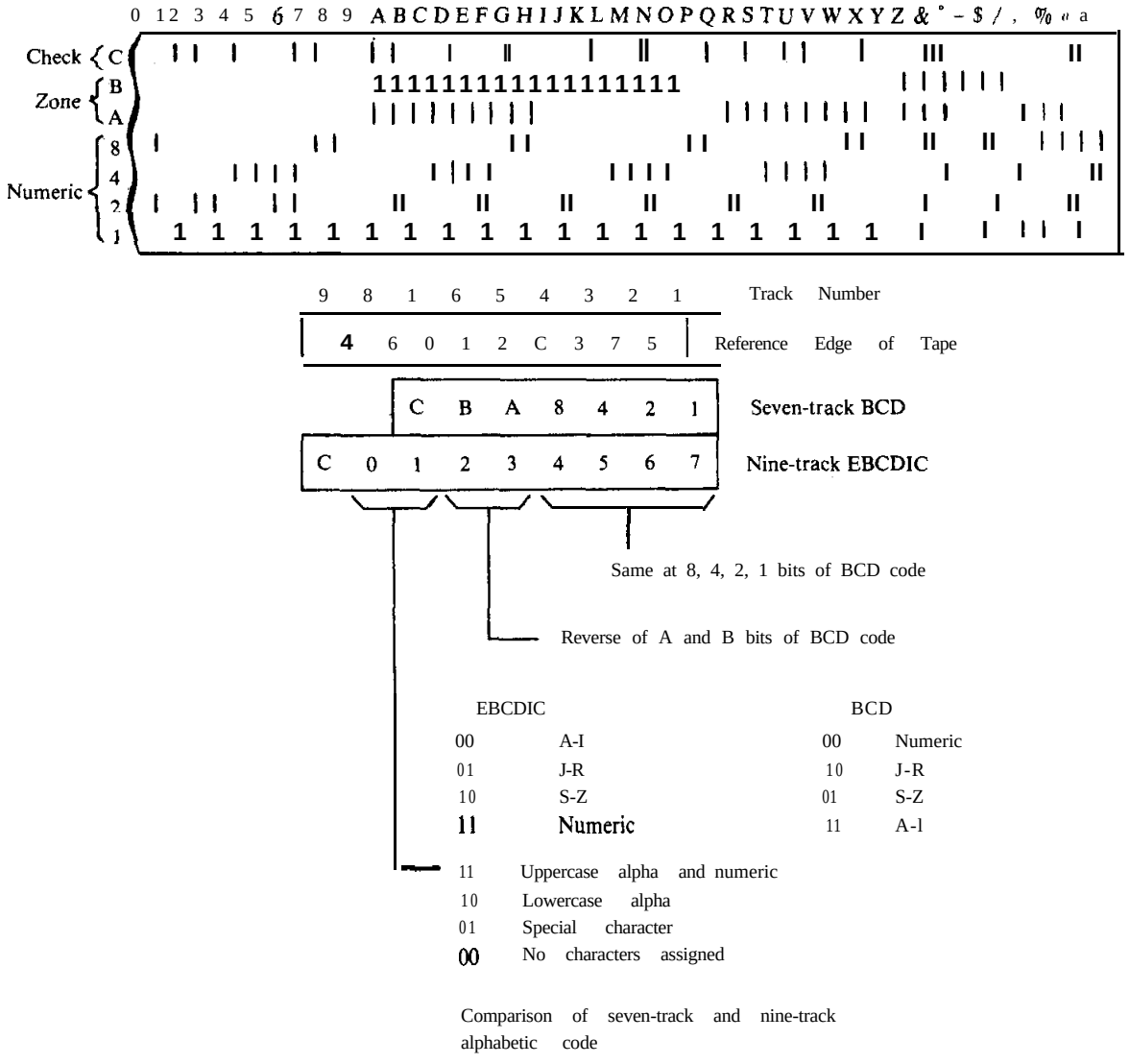
การทำงานของเทปแม่เหล็กเป็นระบบที่เรียกว่า Indirect Access หรือที่เรียกว่า Sequential File Medium เนื่องจากว่าในระบบนี้การอ่านหรือบันทึกข้อมูลต้องเริ่มตั้งต้นตั้งแต่แรกเสมอจึงทำให้เสียเวลา

การบันทึกข้อมูลลงเทปแม่เหล็ก

ดังที่ได้กล่าวมาแล้วแต่ต้นว่าเทปแม่เหล็กเป็นแผ่นยาวบรรจุเป็นม้วน (Reel) นอกจากจะแตกต่างกันในเรื่องความจุ (density) แล้ว เรายังสามารถแบ่งลักษณะของเทปแม่เหล็ก

ออกโดยอาศัยลักษณะที่แตกต่างกันในเรื่องของการบันทึกข้อมูล กล่าวคือ แบ่งตามจำนวน tracks ที่ใช้ในการบันทึกข้อมูล

การบันทึกข้อมูลลงในเทปที่ใช้กันอยู่ทั่วไปมีอยู่ 2 แบบคือชนิด 7 Tracks และชนิด 9 Tracks ชนิด 7 Track ใช้ระบบ BCD (Binary Coded Decimal) และชนิด 9 Tracks ใช้ระบบ EBCDIC (Extended Binary Coded Decimal Interchange Code) การที่จะใช้ระบบใดในการบันทึกนั้นก็ขึ้นอยู่กับชนิดของคอมพิวเตอร์ที่ใช้



Comparison of seven-track and nine-track alphabetic code

หมายเหตุ C จะทำหน้าที่เป็น Check Bit สำหรับตรวจสอบข้อมูลที่จะบันทึก ซึ่งจะกล่าวถึงหน้าที่ของ Check Bit ในบทที่ว่าด้วยเรื่องของระบบเลขจำนวน

เทปแม่เหล็กแต่ละม้วนจะสามารถนำไปใช้ได้หลายครั้ง เช่นเดียวกับเทปบันทึกเสียงเพียงแต่ลบข้อมูลเดิมที่ไม่ต้องการออกแล้วบันทึกข้อมูลใหม่ที่ต้องการลงไปแทน ข้อมูลเก่าก็จะถูกลบออกไปอย่างอัตโนมัติ ผิวข้างหนึ่งของแผ่นเทปแม่เหล็กที่ใช้บันทึกข้อมูลจะเคลือบด้วยสารแม่เหล็ก ดังนั้นเวลาใช้จึงไม่ควรไปจับเนื้อเทป เพราะอาจจะทำให้ข้อมูลที่เก็บไว้เสียหายได้ ดังนั้นที่ต้นม้วนและปลายม้วนเทปซึ่งมีความยาวประมาณ 6 ถึง 8 ฟุต จะเป็นส่วนที่ไม่ใช้บันทึกข้อมูล แต่มีไว้เพื่อให้สามารถจับต้องได้เวลาจะใส่เข้าตู้อ่าน บันทึกข้อมูล (ตู้เทป)

การใช้เทปแม่เหล็กจะต้องใช้กับตู้เทป โดยที่ตู้เทปจะมีหัวอ่าน-บันทึก (Read-Write Head) ตู้เทปนี้จะต้องประกอบด้วย เทป 2 ม้วนคือ File Reel เป็นม้วนเทปที่มีเทปอยู่จะอยู่ทางด้านซ้ายมือของตู้เทป เทปม้วนนี้จะใช้เพื่ออ่านเอาข้อมูลออกมาหรือบันทึกข้อมูลที่ต้องการลงไป

Take Up Reel (Machine Reel) เป็นม้วนเทปที่อยู่ทางขวามือ เป็นม้วนเปล่าไม่มีเทป มีหน้าที่ตอนเก็บเทปจาก File Reel ที่ผ่านเข้าหลังจากอ่านหรือบันทึกจากหัวอ่าน-บันทึกแล้ว เพื่อเป็นการป้องกันความผิดพลาดที่เกิดขึ้น ในขณะที่เราจะบันทึกข้อมูลลงในเทป ต้องใส่อุปกรณ์ชิ้นหนึ่งมีลักษณะเป็นพลาสติกกลม ซึ่งเรียกว่า File Protection Reel เข้ากับม้วนเทปเสมอ โดยที่วงแหวนดังกล่าวจะไปกดกับสวิทช์ที่ตู้คอมพิวเตอร์นี้จะทำให้สามารถบันทึกข้อมูลลงม้วนเทป แต่เราต้องการให้อ่านข้อมูลออกมาก็ถอดวงแหวนดังกล่าวออกเสีย ทั้งนี้เพื่อป้องกันความผิดพลาดอันอาจเกิดจากการบันทึกข้อมูลอื่นลงในข้อมูลเดิมที่ต้องการอยู่อีกได้

การบันทึกข้อมูลหรืออ่านข้อมูลออกจากม้วนนั้นจะทำกันทีละ Block ทั้งนี้เพื่อให้เครื่องคอมพิวเตอร์สามารถนำข้อมูลที่อ่านมาแล้วไปประมวลผลได้เพื่อให้เครื่องคอมพิวเตอร์บันทึกหรืออ่านไปทีละช่วง แล้วหยุดก่อนที่จะไปทำกับข้อมูลชุดต่อ ๆ ไป สาเหตุที่หยุดเพราะว่าที่เทปแม่เหล็กจะมีช่องว่างระหว่าง Block Inter Record Gap IRG โดยที่ IRG นี้ยาวประมาณ $\frac{1}{4}$ - 1 นิ้ว Block แต่ละ Block ดังกล่าวจะเป็นที่เก็บข้อมูล จำนวนข้อมูลที่เก็บใน Block จะวัดกันเป็น Record เช่น 1 Block เก็บได้ 5 Records หรือ 1 Block เก็บได้ 10 Record เป็นต้น สุดแต่ความยาวของข้อมูลแต่ละ Records

เพื่อให้เข้าใจคำว่า Record จะขออธิบายย้อนไปถึงส่วนที่เกี่ยวข้องดังนี้

Bit : บิทเป็นเลขหนึ่งตัวในเลขฐานสองคือ 0 หรือ 1 เรียกว่า 1 บิท ดังนั้น

1011 เรียกว่าเลขจำนวนนี้ประกอบด้วย 4 บิทเป็นต้น

Byte : คือจำนวน Bits ชุดหนึ่งในเลขฐานสองที่ใช้แสดงแทน Character 1 ตัว ใน 1 Byte อาจจะมี 6, 7 หรือ 8 บิทสุดแล้วแต่ระบบเครื่องคอมพิวเตอร์ที่บริษัทผู้ผลิตจะสร้างขึ้น

Character : หรืออักขระในทางคอมพิวเตอร์หมายถึงสัญญลักษณ์ที่ใช้ในระบบดำเนินกรรมวิธีข้อมูลหรือประมวลผลข้อมูลซึ่งได้แก่ตัวอักษร ตัวเลข หรือเครื่องหมายพิเศษต่าง ๆ

Field : เป็นพื้นที่เฉพาะอันหนึ่งที่ใช้สำหรับบันทึกข้อมูลในแต่ละรายการ เช่น ชื่อ นามสกุล ตำแหน่ง เงินเดือน ฯลฯ

Record : คือข้อมูลชุดหนึ่งที่แสดงองค์ประกอบหลักของแฟ้มข้อมูลแฟ้มหนึ่ง หรือแสดงรายการที่เปลี่ยนแปลง (Transaction) หรือหมายถึงกลุ่มของข้อมูลกลุ่มหนึ่งที่เกี่ยวข้องหรือหมายถึงกลุ่มของ Field หลาย ๆ Field ที่ประกอบกันขึ้นเป็นหนึ่ง Record ยกตัวอย่างง่าย ๆ ก็คือ ถ้าเราเก็บประวัติของคนงานเอาไว้ในบัตรเจาะรู โดยกำหนดได้ว่าบัตร 1 ใบจะสามารถเก็บข้อมูลประวัติของคนงานได้ 1 คน นั่นก็หมายความว่าบัตร 1 ใบจะเทียบได้กับ 1 Record หากว่าข้อมูลของคนงาน 1 คนต้องใช้บัตรเจาะรูถึง 2 ใบ ก็หมายความว่า 1 Record เท่ากับบัตร 2 ใบ

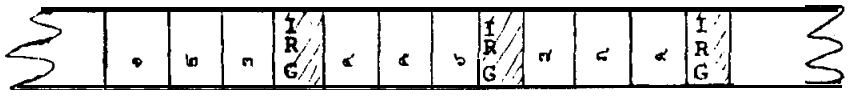
File : แฟ้มข้อมูลคือการรวมและจัดระเบียบ Records ต่าง ๆ ที่มีวัตถุประสงค์ร่วมกันเอาไว้ร่วมกัน เช่นแฟ้มประวัติของคนงานในโรงงานแห่งหนึ่ง แฟ้มประวัติของนักศึกษาคณะวิทยาศาสตร์ มหาวิทยาลัยรามคำแหง เป็นต้น โดยทั่วไปเราจำแนก File ออกเป็น 2 ชนิดคือ

Master File หรือ File ที่เก็บข้อมูลไว้แต่เริ่มทั้งหมด และ File นี้จะถูกเก็บไว้อย่างถาวร แต่ข้อมูลที่เก็บไว้อาจจะเปลี่ยนแปลงได้บ้าง เพื่อให้ถูกต้องตามความเป็นจริง ตัวอย่างเช่น ทะเบียนครัวเรือนในอำเภอหนึ่ง ๆ ประวัติของนักศึกษา (ซึ่งประกอบด้วย ชื่อ ระหัสประจำตัว อายุ ฯลฯ เป็นต้น)

Transaction File คือ File ย่อย ๆ ซึ่งเก็บข้อมูลบางอย่างที่มีความสัมพันธ์กับข้อมูลของ Master File หรืออีกนัยหนึ่ง Transaction File ก็คือ File ที่รวบรวมการเปลี่ยนแปลง

บางอย่างของข้อมูลใน Master File นั่นเอง ซึ่งมักจะทำลายทิ้งไปเมื่อแก้ไข Master File นั่นเอง ให้ทันสมัย (Updating) แล้ว ตัวอย่างเช่น ข้อมูลในการเกิด ตาย โยกย้ายที่อยู่ในระยะเวลาหนึ่งเราถือว่าเป็น Transaction File เมื่อเราแก้ไข (Updating) ทะเบียนบ้าน (Master File) ให้ตรงตามความเป็นจริง

จำนวน Record ที่จะเก็บในแต่ละ Block นั้นเราเรียกว่า Blocking Factor เช่น Block Factor คือ 3 ก็จะได้ว่า



ในกรณีที่ Blocking Factor คือ 1 จะได้ว่า

การที่มี Blocking Factor เท่ากับ 1 บางทีเรียกว่า Unblocked Record และ ถ้ามี Blocking Factor มากกว่า 1 เราเรียกว่า Block Record ในกรณีที่ข้อมูลในแต่ละ Record เป็นชนิดสั้น ๆ และเก็บในเทปแบบ Unblocked Record จะทำให้สิ้นเปลืองเทปไปมาก นอกจากรายละเอียดที่กล่าวมาแล้วนี้ วิธีการเก็บ Record ที่เรียกว่าแบบ Fixed-Length ชนิดที่ 2 ความยาวของแต่ละ Record อาจจะเปลี่ยนแปลงได้ตามจำนวนข้อมูลของแต่ละ Record เรียกว่าแบบ Vareiable Length

Fixed-Length

Variable-Length

แบบความยาวคงที่และความยาวเปลี่ยนแปลงมีข้อดีและข้อเสียแตกต่างกัน กล่าวคือแบบความยาวคงที่ทำให้เปลืองเทปแต่เขียนโปรแกรมใช้งานได้ง่าย ส่วนแบบความยาว