



Enter to grow in wisdom.

*Inscription on the 1890 gate to Harvard Yard*

Since the early days of the abacus, computation devices have been used and sought to make record keeping simpler, to speed computation, and to make our lives easier in general. Developers of computer input equipment are aware of these needs, and their latest equipment is evidence of their user consciousness.

Computers process data at incredible speeds,

but the data must already be in the main storage of the computer for processing to occur. Getting data into the computer is called **input**. It is a slow operation compared to the internal processing speed of the computer. In this chapter we consider the various input devices. We will look at **output techniques** in the next chapter.

A wide variety of input, output, and storage



Data entry  
& display station



Processor & disk storage  
units (to the left)



Card  
read punch



Printer



Tape drive





devices are available. In fact, so many different devices exist that each computer system uses only those few that are appropriate for its applications. These devices are built separately from the processing unit and main storage of the computer, so they are called **peripheral devices**. Figure 4–1 shows some of the input, output, and storage peripheral devices commonly used on larger computer systems.

## Punched Cards

The concept of recording data on cards as a pattern of holes dates from Jacquard's loom of the early 1800s and Hollerith's tabulating machines of the 1880s and is still in use today. Although computer **punched cards** are somewhat outmoded, the phrase, "Do not bend, fold, spindle, or mutilate," which is associated with them, has become a part of our culture.

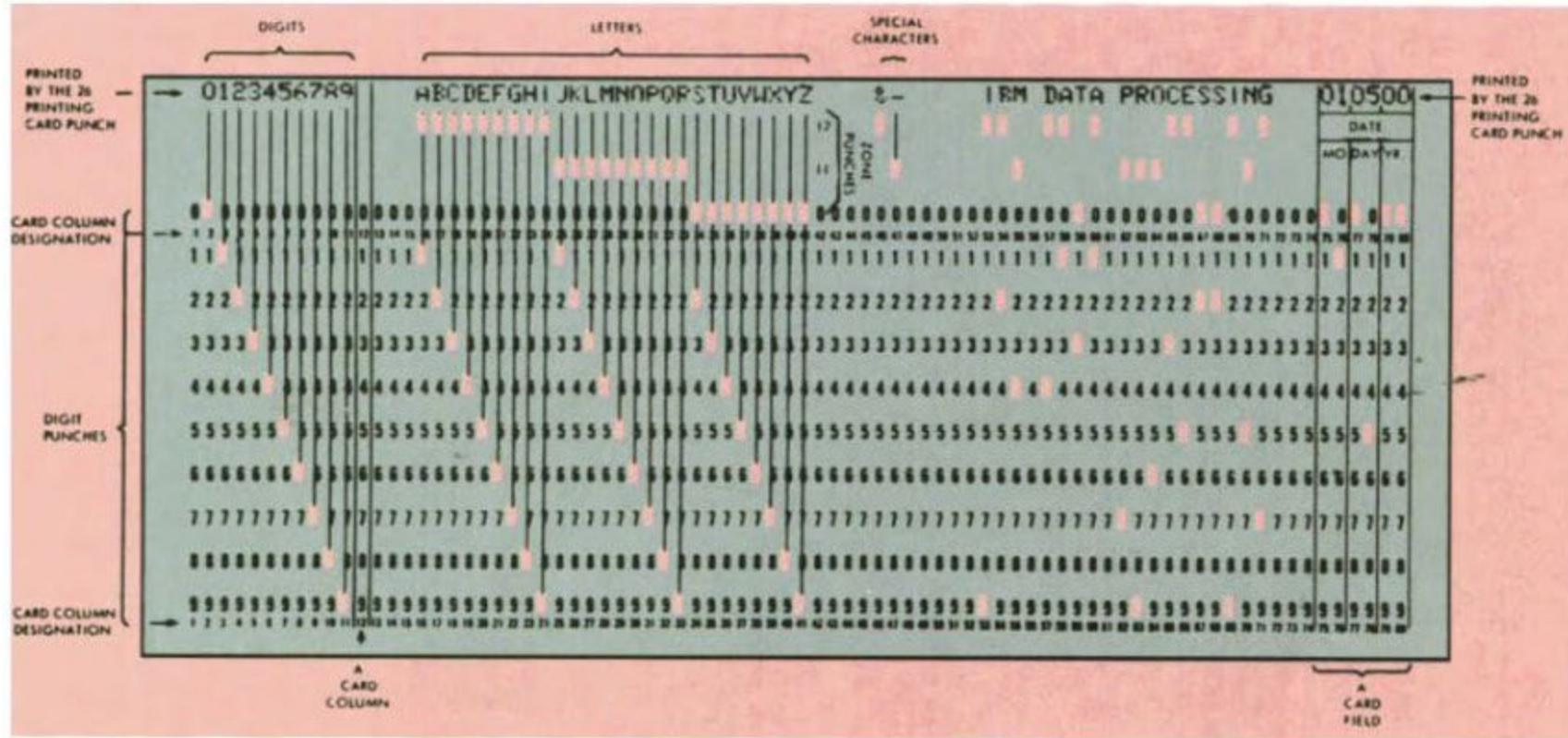
The most common card format is the 80-column card with rectangular holes introduced by IBM in 1928 (Figure 4–2). Each column of the card contains a pattern of holes to represent one

**Figure 4–2** A typical 80-column card showing how digits, letters, and special characters are represented by combinations of holes.

## Garbage-In-Garbage-Out

If you don't put correct information into the computer, you can't produce correct results. *GIGO*, *garbage-in-garbage-out*, summarizes the rule. A famous anecdote about GIGO dates from the early space program days of the 1960s. One of the Gemini (two-person) capsules was due to land at a specific point in the ocean, and all the rescue ships were in place. The capsule's retro-rockets were fired at the exact moment indicated by the computer controlling the flight. Everyone in the rescue fleet looked skyward in great anticipation. TV cameras were set to record the historic moment when the capsule was to appear—but there was no capsule. Great concern followed. When radio contact was re-established with the astronauts and radar tracking picked up the capsule, it appeared more than 100 miles off target! The astronauts landed in the ocean and bobbed around in the waves for several hours, getting quite seasick.

NASA checked the computer equipment and found that it was functioning perfectly. But when the computer program that ordered the retro-rockets to fire was examined, it was found that a piece of bad data had been input. The program had required the number of hours in the day to make its calculations. The programmer had naively entered 24 hours, but a day is not exactly 24 hours in length (that's why we have leap years). The difference accounted precisely for the 100-mile error in the landing. A testament to GIGO!







character, so up to 80 characters of data may be punched in an 80-column card. The card has 12 rows. The various letters, digits, and special symbols are represented by punching one, two, or three holes in a particular column (Figure 4–3).

The large size of the 80-column format worried many users because of the occasional paper shortages and projected increases in consumption. In 1970 IBM announced its System/3 computer system, and with it a new smaller card format. The new cards (Figure 4–4) have space for punching 96 characters in three rows, or tiers, of 32 characters each. They have a number of advantages over the 80-column cards, including being smaller, cheaper, easier to handle, and less costly to store. Nevertheless, the 96-column card has, for the most part, disappeared, for reasons that will soon become clear.

**Figure 4–3** The hole punches that correspond to the digits, letters, and special symbols most commonly used with punched card inputs.

Digits	Holes Punched	Letters	Holes Punched	Symbols	Holes Punched
0	0	A	12–1	&	12
1	1	B	12–2	€	12–2–8
2	2	C	12–3	.	12–3–8
3	3	D	12–4	<	12–4–8
4	4	E	12–5	(	12–5–8
5	5	F	12–6	+	12–6–8

## The Keypunch Machine

Data from a **source document** is punched onto cards by a **keypunch** machine (Figure 4–5). The machine has a typewriterlike keyboard, a hole-punching mechanism, and a card transport mechanism for moving cards through the machine. A stack of blank cards is placed in the machine. One card at a time is moved to the *punching station*. An operator then types the data to appear on the card.

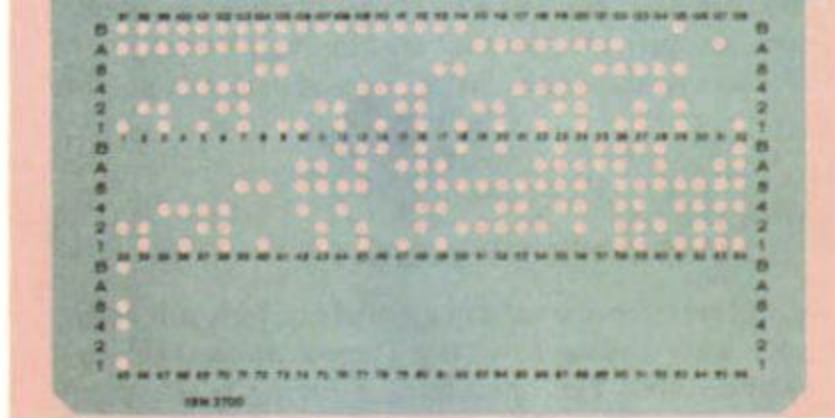
**Figure 4–4** A 96-column card.



**Figure 4-3** The hole punches that correspond to the digits, letters, and special symbols most commonly used with punched card inputs.

Digits	Holes Punched	Letters	Holes Punched	Symbols	Holes Punched
0	0	A	12-1	&	12
1	1	B	12-2	€	12-2-8
2	2	C	12-3	.	12-3-8
3	3	D	12-4	<	12-4-8
4	4	E	12-5	(	12-5-8
5	5	F	12-6	+	12-6-8
6	6	G	12-7		12-7-8
7	7	H	12-8	-	11
8	8	I	12-9	!	11-2-8
9	9	J	11-1	\$	11-3-8
		K	11-2	*	11-4-8
		L	11-3	)	11-5-8
		M	11-4	:	11-6-8
		N	11-5	□	11-7-8
		O	11-6	/	0-1
		P	11-7	,	0-3-8
		Q	11-8	%	0-4-8
		R	11-9	—	0-5-8
		S	0-2	>	0-6-8
		T	0-3	?	0-7-8
		U	0-4	∴	2-8
		V	0-5	#	3-8
		W	0-6	(	4-8
		X	0-7	'	5-8
		Y	0-8	=	6-8
		Z	0-9	"	7-8

(Blank)



Courtesy of International Business Machines



"Murder, huh? I'm in for bending, folding and mutilating computer cards."





**Figure 4-5** The keypunch machine has been one of the most popular data entry devices of all time.

## Card Verification

To ensure that the keypunched data is correct before it enters the computer, a **card verifier** may be used. A verifying operator feeds in the cards punched by the keypunch operator and retypes all the information. The card verifying machine does not punch holes; it merely checks that the holes already punched in the card are the same as the characters being retyped by the card verifying operator. If they are, the card is considered correct, and it may be entered into the computer. If

1. **Edit checks** ensure that digits are not entered where letters are required, and vice versa. If the edit program discovers such an error, the card is listed on a printed report with an indication of the error. A person may then correct and resubmit the card.
2. **Range checks** ensure that a data item has a value within a specific range. For example, a date should contain a month number from 1 to 12, but months of 0 or 14 would be considered errors.
3. **Reasonableness checks** ensure that a data item is reasonable. Data items that fail reasonableness checks are not always incorrect. For example, a reasonableness check on a person's age could be "less than 150," but it is possible, although unlikely, that a person would live to that age.

## Reading Cards into the Computer

Once cards have been punched and verified, they are entered into the main computer system

holes already punched in the card are the same as the characters being retyped by the card verifying operator. If they are, the card is considered correct, and it may be entered into the computer. If they are not, the verifier flashes a warning, the card is marked for repunching, and a corrected card is punched and inserted in the card deck to replace the bad card.

Verification is an expensive and time-consuming process. It ensures high reliability, but it does not guarantee perfection. The verifying operator should not be the same person as the keypunch operator, or the same mistakes might be made in verification as were made in keypunching. Verification is expensive, but it is often cheaper than correcting errors after they have entered the computer system.

## Eliminating Verification

The risk of errors with manual keystroking is great. Techniques have been developed, however, to eliminate verification without increasing input error rates. Most of these methods use the computer system itself to scan incoming cards for certain common errors. By using the following techniques, many organizations have eliminated the need for verification.

Once cards have been punched and verified, they are entered into the main computer system through a device called a **card reader**. The computer operator places the deck of cards in the **input hopper**. Cards are then read one at a time as they pass a **read station** and are then stacked in an **output hopper**. One popular reading technique is to shine intense light at the card. Where a hole is punched, the light shines through; where there is no hole, the light is blocked. Photosensors detect light passing through the holes and the computer reads the photosensors, converting the hole patterns into corresponding bit patterns. Some high-speed card readers can process more than 1000 cards per minute.

The card reader is “affectionately” called the card chewer by many frustrated operators. Handling paper at high speeds is a difficult task, and occasionally an operator has helplessly stood by watching a malfunctioning card reader maul a deck.

## An Assessment

Punched cards were the primary data entry medium through the early 1970s when, in the United



States alone, about half a million keypunch machines were in use, and most computer systems had one or more high-speed card readers. Their use is declining rapidly now because of their many disadvantages. Cards are bulky to handle and expensive to produce, process, and store. They cannot be reused and can file only as much information as can be stored in 80 characters (the more popular tapes and disks can store thousands of times as much information in the same amount of space). **File integrity** problems are possible with cards; they can be removed from files, replaced with other cards with altered data, and easily lost or stolen. Because of the availability of improved input methods, the vast majority of computer systems built today do not even have card readers.

## Key-to-Tape and Key-to-Disk Systems

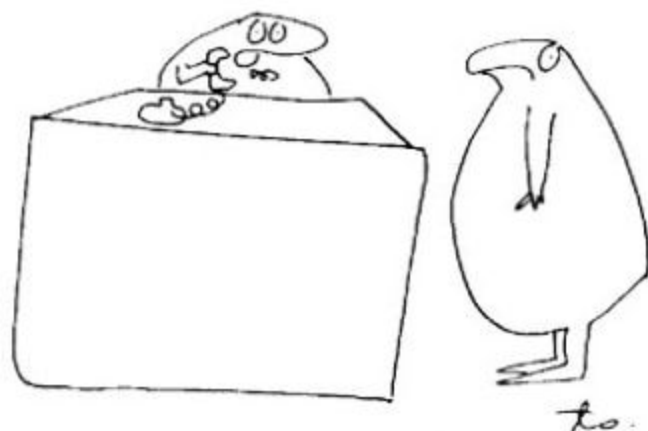
Because of the many disadvantages of punched cards and the continuing decline in the price of computer power, it became economical in the early 1970s for larger data entry installations to

Although the systems are more complex than keypunching, they are more economical.

## Terminal-Oriented Systems

A large portion of users now operate **on-line** to computers, that is, through terminals or keyboards directly connected to computers or data communications networks. Timesharing (see Chapter 12 for a discussion of timesharing) has made it possible for hundreds and even thousands of users to be on-line to a single computer system

©INFOSYSTEMS



"I guess our customer training wasn't too good. The company wants to install a keyboard-to-wastebasket system."

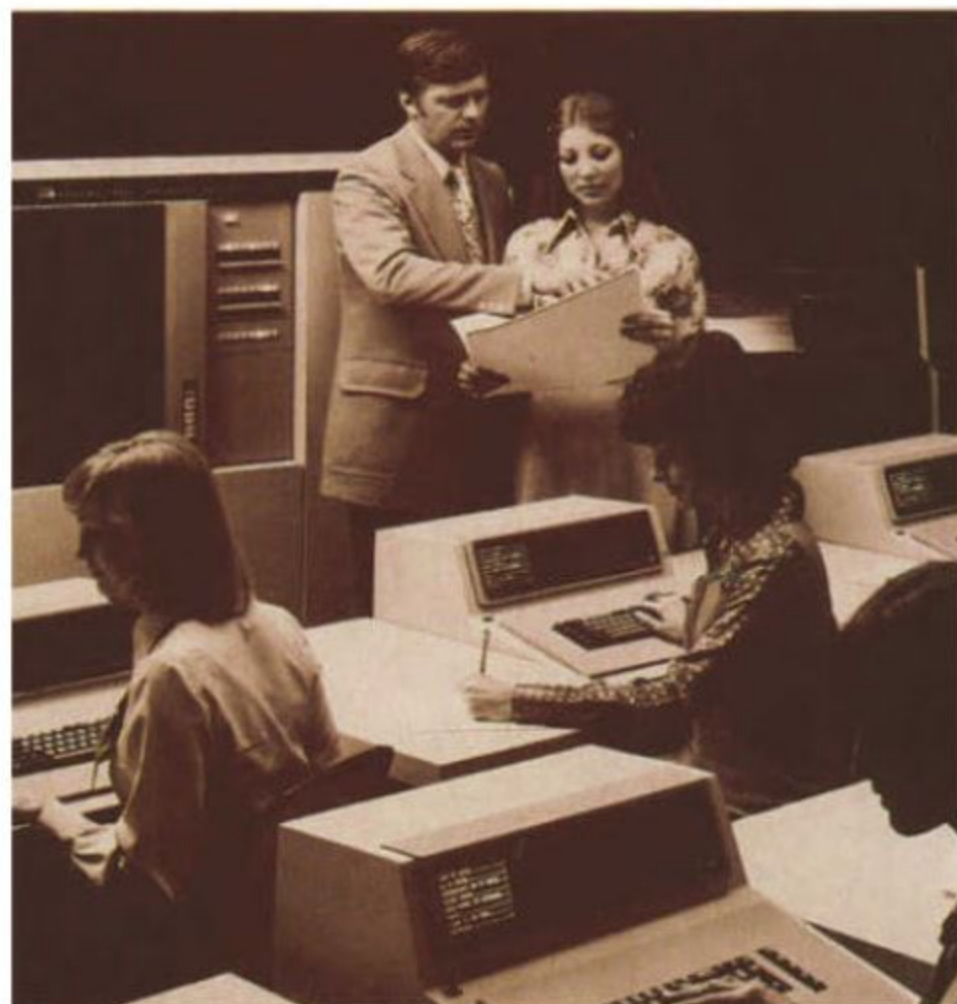
Because of the many disadvantages of punched cards and the continuing decline in the price of computer power, it became economical in the early 1970s for larger data entry installations to replace keypunch machines with **key-to-tape** (Figure 4-6) and **key-to-disk** (Figure 4-7) machines. These are sometimes called **intelligent data entry devices** because they have their own built-in computers to control the data entry process. Instead of punching holes in nonreusable cards, operators type on a keyboard to enter computerized data directly onto reusable magnetic tapes and disks.

**Figure 4-6** Key-to-tape devices generally have only a single data entry station.



to.  
"I guess our customer training wasn't too good. The company wants to install a keyboard-to-wastebasket system."

**Figure 4-7** Key-to-disk devices generally handle many data entry stations simultaneously.







at once. Personal computers generally accept their inputs directly from attached keyboards.

This on-line operation is another improvement over punched cards. With punched cards users operate **off-line** from their computers. They submit work and receive results through devices not directly connected to the computers. This method of operation is often called **batch processing**, in which data is collected, entered into the computer, and processed in groups called **batches**. Results may be returned hours later.

On-line, terminal-oriented systems are often called **transaction-processing systems**. Individual data items are entered into the computer and processed immediately—each of these items is called a transaction.

## Touch Sensing

Over the years many people have viewed computers as rather unfriendly devices and have been re-

luctant to work with them. One way the industry is making the user/computer interface more user-friendly to the general public is through the technology of **touch sensing**. An operator does not need to be an expert typist and does not have to be intimately familiar with the often messy details of running a computer. Instead, the operator merely points to drawings, diagrams, symbols, or words on a screen (Figure 4–8). The computer senses that a particular area of the screen is being touched and responds accordingly. People seem to enjoy pointing to what they want. In fact, one study estimates that 95 percent of the population could competently and easily operate a touch sensitive device without prior training.

For touch sensing to be effective, the number of choices displayed on the screen at one time should be small, perhaps 10 or fewer. Having too many choices may lead to confusion. Because human fingers are large compared to individual characters on a screen, it is impractical to display a large number of choices anyway.

With touch sensing, any symbol, picture, or customized set of buttons useful for a particular

Over the years many people have viewed computers as rather unfriendly devices and have been re-

**Figure 4-8** The touch control screen of the Xerox 5700 Electronic Printing System gives the operator control over all system operations. The screen replaces the usual operator control panel found on similar devices. The screen displays pictures of buttons; the operator touches the screen where a button is displayed to tell the machine to perform the function indicated by the button.



ingers are large compared to individual characters on a screen, it is impractical to display a large number of choices anyway.

With touch sensing, any symbol, picture, or customized set of buttons useful for a particular application may be displayed. When a regular typewriter keyboard is appropriate, a conventional keyboard may be displayed. When a numeric keyboard is appropriate, a 10-key numeric pad can be shown.

The display can be tailored to the needs of the application and to the educational level of the users. For example, a computer-assisted instruction (CAI) system for kindergarten children might contain simple pictures, whereas one for older

## Touch Tour

The concept of touch-sensing terminals has been adopted by the developers of Walt Disney's EPCOT (Experimental Prototype Community of Tomorrow) located near Orlando, Florida. Scattered throughout EPCOT are information stations that use touch sensing. Once a visitor indicates the main topic, the computer displays a menu, or list of subtopics. By pressing numbers or letters, visitors are guided along until the information they want appears on the screen. Besides being informative, the system is fun to use, even for people with no computer training.





students might use a substantial vocabulary of words and phrases. Engineers could use a system with symbols for electrical components, and football coaches could use one with symbols for each of the key players.

Figure 4–9 shows many ways of inputting touch-sensitive data on a **cathode ray tube (CRT)**, a display unit attached to a computer terminal. In diagram (a), the user touches the screen directly (also see Figures 4–11 and 4–13). In Figure 4–9, diagram (b), the user “writes” on the screen with a device called a **light pen** (also see Figures 4–10 and 4–14). The light pen detects light from the characters on the screen and then sends a message to the computer indicating the position the user is pointing to. In diagram (c) of Figure 4–9, the user writes on a data tablet, which senses the

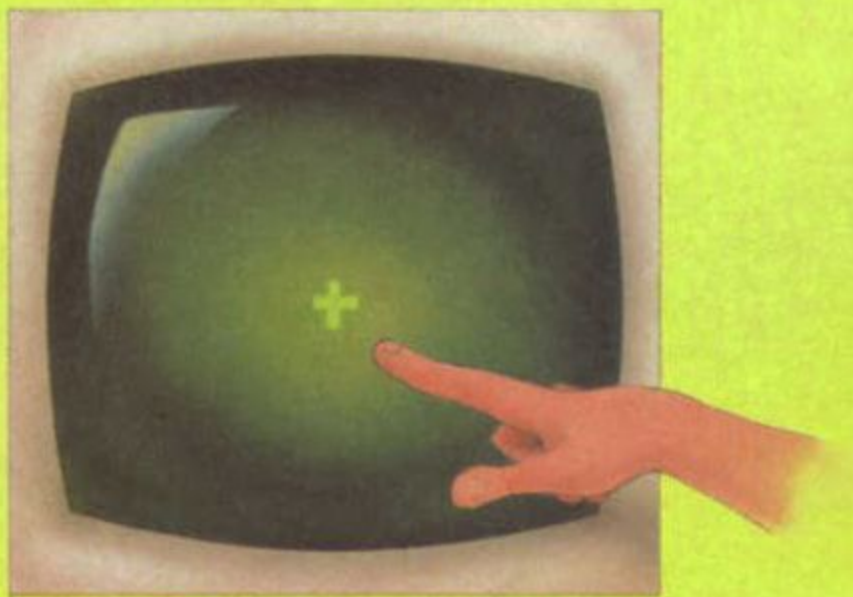


Figure 4–10 A light pen.

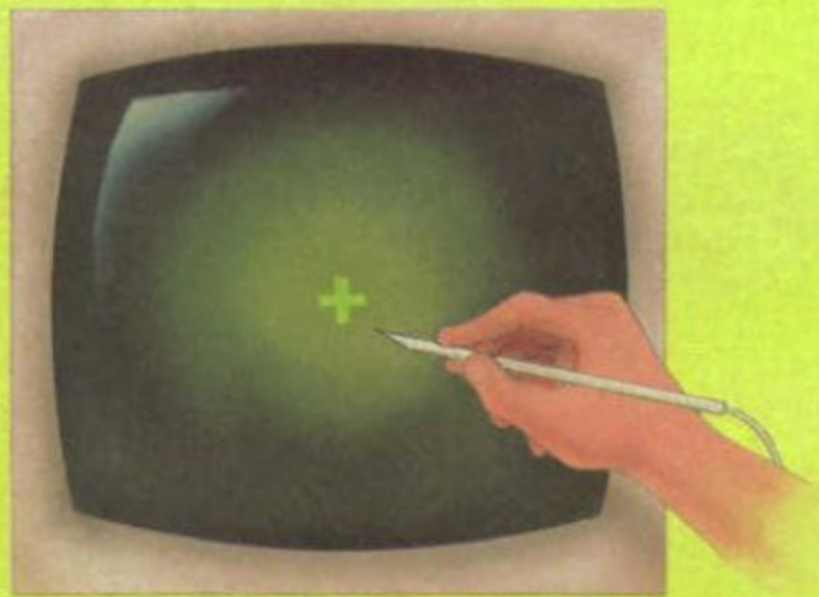
Figure 4–9 Inputting touch-sensitive data.



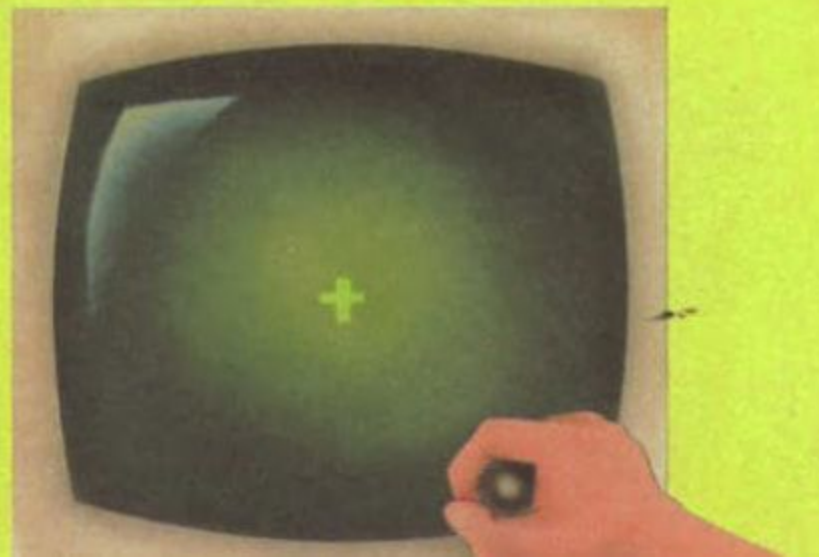
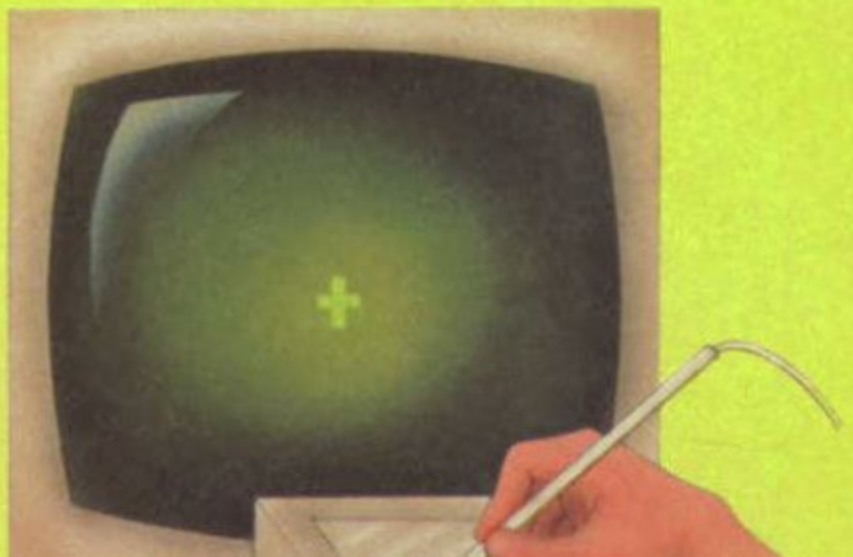
**Figure 4-9** Inputting touch-sensitive data.



**(a) Direct Tactile**



**(b) Direct Graphic**







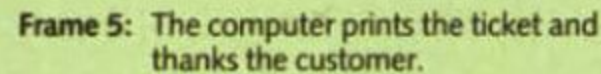
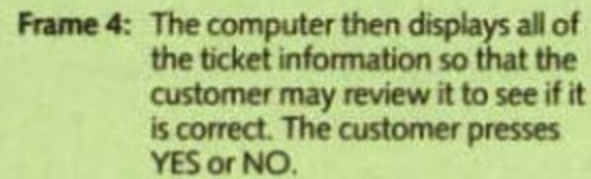
**Frame 1:** The passenger presses the screen to select DENVER as the destination city.



**Frame 2:** The computer displays today's flights to Denver. The person Touches the screen on the desired flight number, and then touches either the word COACH or the phrase FIRST CLASS.



**Frame 3:** For seat selection, the passenger then presses WINDOW, CENTER or AISLE. The section is then selected — either SMOKING or NON-SMOKING.



**Figure 4–11** Here a touch sensitive screen is used in a system designed for airline passengers to help them make their own reservations and write their own air tickets automatically.





position of the pen and causes the figure being drawn by the user to appear simultaneously on the screen (also see Figure 4–12). In Figure 4–9, diagram (d), the user manipulates a **joystick** like that used with video games; as the joystick is moved, the cursor moves in the corresponding direction on the screen. (Figure 4–15 shows an input device called a **cursor**.)

CRTs are called **soft copy** devices because they do not produce a paper copy (**hard copy**) of the displayed data. Hard copy can be produced when the CRT is connected to a printer and when the user specifically requests it.

## Other Input Devices

Another type of user-friendly system is the computerized **speech recognition** device. This input device analyzes spoken commands and converts them to computer-stored data. A surgeon can issue commands to control devices in the operating room. A handicapped person can use voice commands to control the movement of a wheelchair.



Figure 4–12 Small digitizer tablets are available for use with personal computers.

room. A handicapped person can use voice commands to control the movement of a wheelchair. Security systems use speech recognition devices to identify people.

A TV camera (Figure 4-16) can also work as an input device. It can be designed to convert a picture into digital data, that is, to digitize the picture. The picture may then be printed on a computer printer.

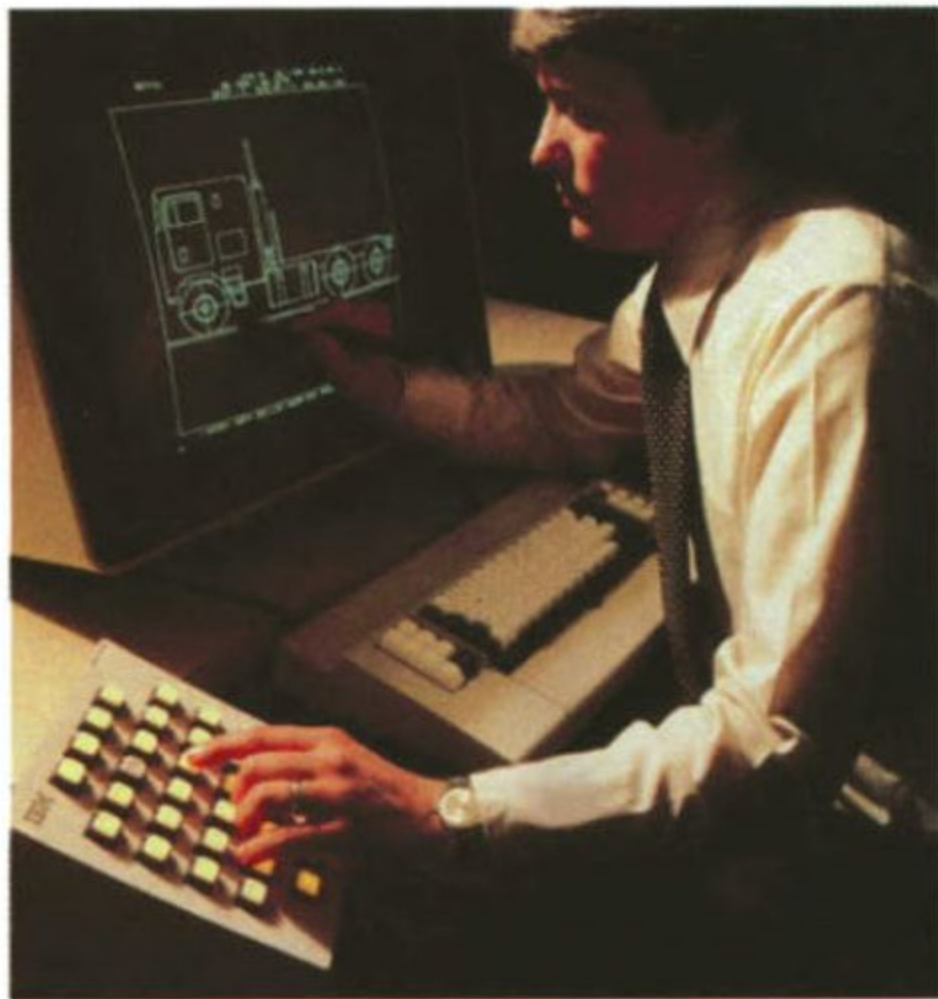
Other input devices use so-called membrane keyboards. The membrane (Figure 4-17) has no movable keys. Instead it is a flat surface imprinted with pictures of the keys. When the picture of a key is pressed, it causes the flat surface to come in contact with another flat "membrane" behind it. An electrical signal is generated to indicate that the key in that position has been pressed. Membrane keyboards cost less than those with moving parts and resist damage by cigarette ashes, spilled liquids, and dust.

Another innovative device is the three-dimensional digitizer (Figure 4-18). The user moves the hand-held stylus around a three-dimensional object (in this case a plane model). Information about the shape of the object is conveyed to the computer, which can then display the object from any angle. Figures 4-19 and 4-20 on page 84 show other interesting input devices.

**Figure 4-13** In this inventory control application, the user may call up one of eight different inventory control functions by pressing the appropriate rectangle on the touch sensitive screen. The bottom rectangle allows the user to "EXIT" from the inventory application; this causes a fresh screen to be displayed with other options.







**Figure 4-14** This person is designing a truck body with the assistance of a CAD (Computer Assisted Design) system. He may modify the design by pointing to the appropriate lines or shapes on the screen with a light pen and then pressing one of the control buttons at the left.



**Figure 4-16** Above left: A TV camera may also be used as a computer input device. This system uses a TV camera to record a person's picture. The computer "digitizes" the picture into a series of lighter and darker colored areas, and prints the picture on a computer printer.

**Figure 4-17** Above right: A membrane keyboard.

**Figure 4–15** This small hand-held device is called a cursor. The user positions it above a word, phrase, or diagram on a digitizing tablet. The user then presses one of the buttons to transmit the cursor position to the computer, thus telling the computer what is being selected by the user.



**Figure 4–18** A three-dimensional digitizer.







**Figure 4-19** This input device uses a pressure-sensitive writing surface to capture handwritten data at the time of writing. Anyone who can write can use this simple device to enter information into a computer.

**Figure 4-20** This input device is called a mouse. The mouse sits on a flat surface next to a computer terminal, and the person using it moves it in a given direction. Roller bearings on the bottom of the mouse detect motion. This information is transmitted to the computer, and a blip displayed on the

## Source Data Automation

Data processing has traditionally begun with the conversion of data documents to computer-readable form by manual data entry techniques such as keypunching. As we have seen, this process is slow, labor intensive, expensive, and prone to error.

Bypassing manual data entry is called **source data automation (SDA)**. The goal of SDA is to capture data automatically and in computer-readable form where it originates. The advantages are obvious—reduced cost, elimination of duplicate effort, reduced need for skilled data entry operators, fast and accurate data entry, and faster results because of the reduced input times.

Several techniques for source data automation are widely used, including: (1) **magnetic ink character recognition (MICR)**, used mostly in automated check processing; (2) **optical character recognition (OCR)**, used in applications where the computer reads human-readable symbols directly; (3) **optical mark recognition (OMR)**, used in exam scoring, survey taking, and utility meter

**Figure 4-21** *Below:* Data may be entered directly into the computer through computer terminals like the one shown here.

**Figure 4-20** This input device is called a mouse. The mouse sits on a flat surface next to a computer terminal, and the person using it moves it in a given direction. Roller bearings on the bottom of the mouse detect motion. This information is transmitted to the computer, and a blip displayed on the screen moves in the same direction as the mouse. So the mouse may be used to point to items on the screen.



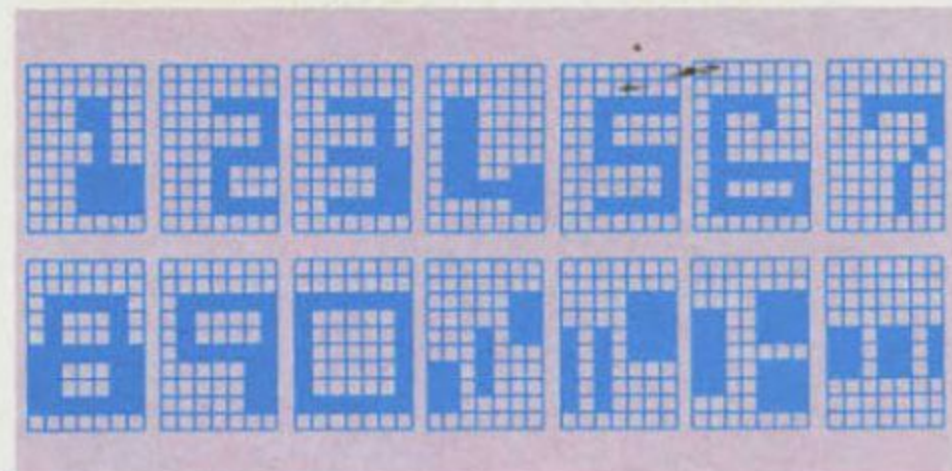
in exam scoring, survey taking, and utility meter

**Figure 4-21** *Below:* Data may be entered directly into the computer through computer terminals like the one shown here.

**Figure 4-22** *Bottom:* The grid pattern used by an MICR reader to help recognize characters.



Courtesy of International Business  
Machines Corporation



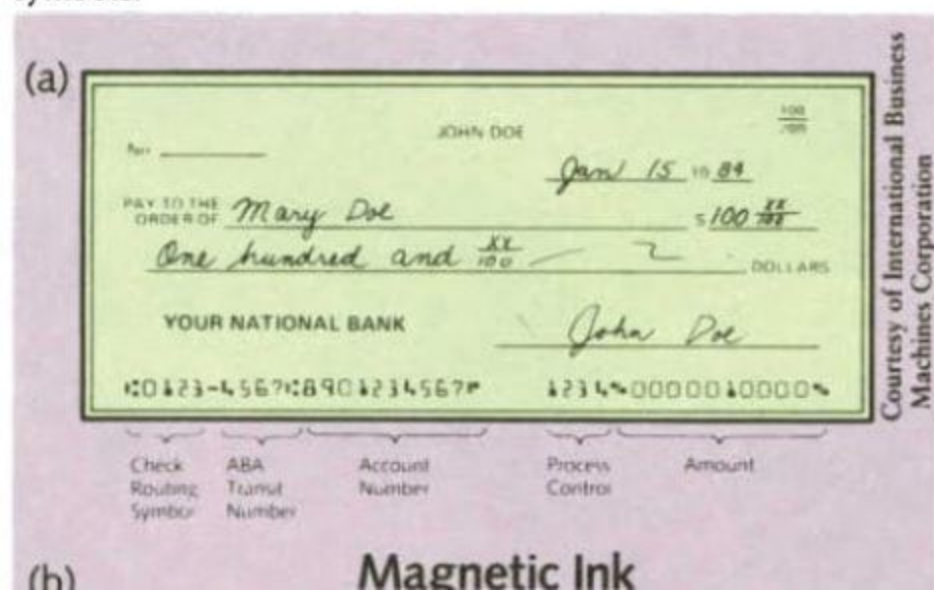
Courtesy of International Business  
Machines Corporation



reading applications; and (4) bar code reading, used in inventory systems and in automated supermarket checkout systems.

The trend is moving away from even these popular techniques. In fact, most input will eventually be performed by **direct data entry**, where the data is captured at its source and entered directly into the computer with no paper documents being produced (Figure 4-21). Direct data entry is now used, for example, when an airline reservationist types information about a flight into the computer while speaking with a customer on the telephone.

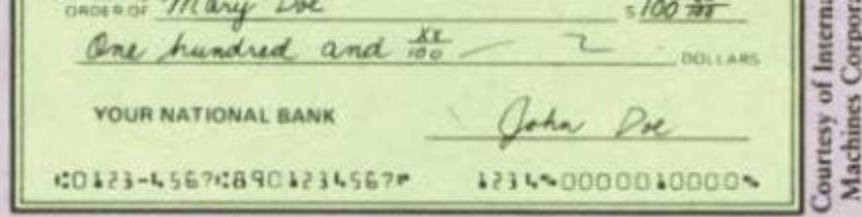
**Figure 4-23** (a) A sample check using MICR characters. (b) The commonly used MICR characters and special symbols.



## Magnetic Ink Character Recognition

Magnetic ink character recognition is a technique developed by the banking industry to speed check processing. The ink used in MICR characters is magnetic, and MICR readers are sensitive to the magnetized ink. The sensors in an MICR reader recognize each character by comparing it to a grid pattern (Figure 4-22). Figure 4-23 shows a sample check with MICR characters encoded across the bottom. The numbers at the bottom left of the check identify the bank and the account number. The amount of the check at the bottom right is typed manually in the process of clearing the check. Many checks also contain the check number printed in MICR characters along the bottom, which makes it possible to identify checks by number on a customer's computer-prepared bank statement. The MICR character set contains the digits 0 to 9 and four special-purpose symbols used to separate and identify the pieces of information on the bottom of a check.

Many large banks process millions of checks each day. To handle such volumes, they use high-speed MICR readers, each capable of processing 50,000 to 100,000 checks per hour. This means that the reading mechanisms (Figure 4-24) have to handle as many as 20 or 30 checks per second



Courtesy of International Business Machines Corporation

(b)

## Magnetic Ink Character Recognition Chart

0	1	2	3	4
Zero	One	Two	Three	Four

5	6	7	8	9
Five	Six	Seven	Eight	Nine

	Amount symbol
	On US symbol
	Transit number symbol
	Dash symbol

purpose symbols used to separate and identify the pieces of information on the bottom of a check.

Many large banks process millions of checks each day. To handle such volumes, they use high-speed MICR readers, each capable of processing 50,000 to 100,000 checks per hour. This means that the reading mechanisms (Figure 4-24) have to handle as many as 20 or 30 checks per second accurately and without damaging them. If that doesn't sound impressive enough, consider that both sides of a check are microfilmed as it passes through an MICR reader.

**Figure 4-24** A check passing through an MICR reader. During processing the paper moves so quickly through this device that it can't be seen.



Courtesy of International Business Machines Corporation



## Optical Character Recognition

Optical character recognition devices read handwritten or typewritten characters and symbols and convert them directly into computer codes faster than people can type and at far less cost.

OCRs can be classified according to their reading capabilities. These capabilities are described in terms of a font, that is, the complete assortment of letters, digits, and special characters of one style and size. **Hand-print readers** can read hand-printed characters. The characters must be printed very carefully in order to be readable (Figure 4–25).

**Printed-font readers** read typewritten and type-set characters. They are classified according to the variety of fonts they read:

1. **Single-font readers** are normally used in special-purpose applications such as reading gasoline purchase slips with credit card impressions.
2. **Multifont readers** are used in data-processing operations that must service many different types of OCR applications.








































3. **Omnifont readers** are the most versatile and most expensive devices on the market today. They can read most printed fonts and are often equipped with complex programs that help them learn new fonts. The premier machine of this type is the Kurzweil OCR reader (see the box on page 89).

Some of the more popular OCR fonts are shown in Figure 4–26. The most widely used OCR font in the United States is **OCR-A**. Many organizations are pressing to change to **OCR-B**, which is used internationally because most people find it easier to read, but studies have shown that OCR-A is recognized by reading machines far more reliably than OCR-B.

OCR devices have been built to read a variety of documents. These readers fall into the following classes:

1. **Journal tape readers** are designed primarily for the retail trades. These devices read the narrow paper printouts produced by cash registers and adding machines.

**Figure 4–25** Rules for correct printing of hand-printed characters in OCR applications.

Rule	Right	Wrong
Characters should be in boxes	  	  
Character should almost fill the box	  	  
Loops should be closed	  	  
Characters should not connect	  	  
Adjacent lines should connect	  	  
Shapes should be simple	  	  
Block-style characters should be used	  	  



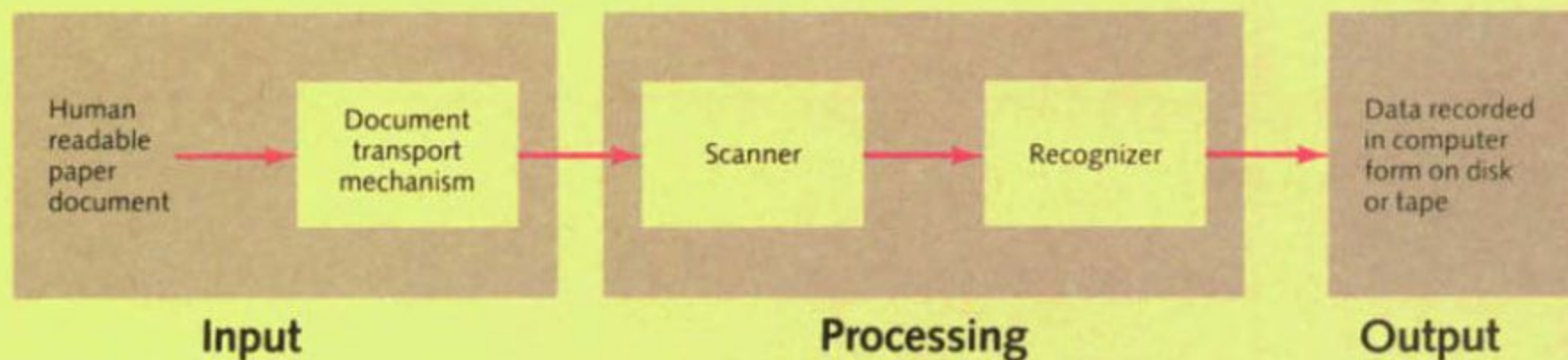


Figure 4–27 Organization of a typical OCR reader.

2. **Small document readers** are used primarily by the utility and gasoline retailing industries. These devices read the credit card imprinted stubs from gasoline stations, meter readings, and bills returned with payments.
3. **Page readers** read full pages of text, and some are capable of reading a variety of fonts.
4. **Wand readers** are small hand-held units that bring OCR reading capability directly to the source of the data. They are widely used in the retail trades.

## How OCR Works

Figure 4–27 shows the organization of a typical OCR reader. Documents are placed in an input

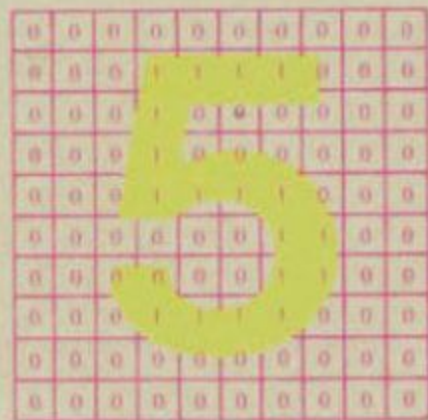
nized as a vertical line with a dot above it; O is recognized as a closed curve without any straight portions that might, for example, confuse it with the letter D. Some OCRs place documents with unrecognizable characters in a reject stacker. Others display unrecognizable characters on a screen, and the operator reenters them at a keyboard.

Optical character recognition is not, and never can be, 100 percent accurate in reading hand-printed characters. Consider, for example, the reading of hand-printed catalog codes known to contain both letters and digits. People frequently confuse the numeric zero with the letter O, the letter S with the number 5, and the letter Z with the number 2. So OCR reading of hand printing

## How OCR Works

Figure 4–27 shows the organization of a typical OCR reader. Documents are placed in an input stacker and transported to a scanning mechanism. Light is reflected off the page onto light-sensitive photocells that register the presence or absence of light. The photocells are “digitized” into a pattern of bits (Figure 4–28). Computer programs analyze these patterns looking for the lines, curves, and dots that form the various symbols to be read. For example, the letter *i* is recog-

**Figure 4–28** The scanner reads the “5” by converting it to a bit pattern. The recognizer compares the bit pattern to a set of prestored patterns to see if there is a match.



ning of hand-printed characters and will contain both letters and digits. People frequently confuse the numeric zero with the letter O, the letter S with the number 5, and the letter Z with the number 2. So OCR reading of hand printing should not be used in situations where very high accuracy is critical. Nevertheless, many OCR systems today read hand-printed characters with better than 99 percent accuracy.

## OCR Applications

To speed processing of retail merchandise, many companies place preprinted OCR readable tags

**Figure 4–29** This device is widely used in the retail clothing industry. It prints price tags with OCR readable characters, and pins the tags to garments.

