introduction to data communication

by

donald e. murphy
with
stephen a. kallis, jr.

DIGITAL EQUIPMENT CORPORATION

contents

ILLUSTRATIONS

Figure

preface

Introduction to Data Communication has been written to outline the basic principles of data communication for a reader who has little or no knowledge of this field. To ensure that a maximum number of people can use this text profitably, the first two chapters discuss basic concepts common to all phases of communication, such as communication theory, baud rate, simplex systems, and so forth. Those familiar with these concepts can ignore or skim these two chapters and begin with Chapter 3, which discusses common carrier communication facilities.

foreward

Communicating has always been an important human activity. It has been responsible for the development of cultures, and their evolution or downfall. Until recently, human communication techniques have been limited to the spoken and written word, sign language, and forms of artwork.

During the industrial revolution, new communication techniques began to evolve. Telegraphs, telephones, phonographs, radios, photographs, motion pictures, and television have all been developed in a relatively brief span of time. Communication equipment and techniques are still being developed and improved.

As these communication systems are developed, many of them have become very specialized. As an example, for the relatively limited information required by a stock broker, a few symbols will give him all the information he needs.

Data Communication, although on the surface all encompassing, is really a large but specialized subdivision of communications in general. It is the intent of this handbook to provide an introduction to this increasingly important field.

chapter 1

communication theory

Although different methods of communicating have been studied for many years, it has only been recently that the process of communicating itself has come under study.

The study of the communication process has led to a new discipline, known as *communication theory*. Pioneer work in communication theory was performed by Claude E. Shannon. In his paper "A Mathematical Theory of Communication," he laid the groundwork for much current communication theory, brought together many diverse problems of communication, and established unifying concepts for them.

The minimum elements in any communication process are a message source, a message medium, and a receiver. The simplest case might be one person talking to another. The person talking is the transmitter or message source, air is the separating message medium,

Minimum elements of communication process and the listener is the receiver. Similarly, the authors are the transmitter or message source for this book, the publication is the medium, and the reader is the receiver. This same generalization is true of all communication techniques, whether the message is sent via a phonograph record, letter, semaphore, or photograph. It allows us to see immediately that without each of these minimum communication elements, communication is impossible.

To communicate, information of some sort must be transferred. "Information," in communication theory is defined as any organized signal. It can be a series of letters, a group of tones, a series of picture scanning elements, or the wail or a siren. This information comprises the message.

In order for a communication to be effective, it must be understood. In the example of a minimum communication system, the message is sent one way to a receiver without thought of understanding on the part of the receiver. By this, we are not considering a receiver that cannot detect the message; we are considering a situation where a receiver needs more information to make a message meaningful. For instance, if a person mispronounces or mumbles a word in a conversation, a listener will generally ask him to repeat what he has said, which permits clarification of his message. On the other hand if there is a confusing misprint in a book, the reader may have no method of determining what the author was trying to say, short of hunting down the author to ask him what word he meant.

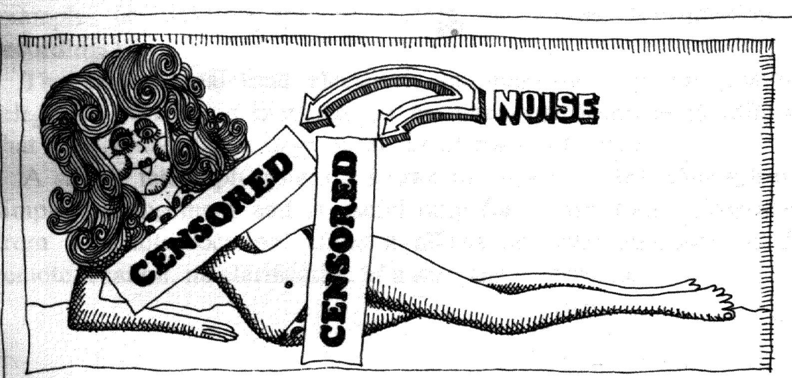
In order for a communication to be effective, it must be understood

A communication system that does not allow an interchange between the message source and the receiver is said to be a one-way or "simplex" communication setup. One that does allow an interchange is a two-way or "duplex" setup.

The medium carrying a message is limited. For example, this book has a limited number of pages, each of which is a standard size. The medium within the limits in which the message is carried is known as the channel. For a private intercom system, the communication channel may be an electrical wire and for a radio station, a range of frequencies.

In a medium like radio or television, a channel covers only a limited range of frequencies to transmit a message. The difference between the upper and lower frequencies (in Hertz) of the channel determines the bandwidth. The nominal frequency of the channel is the frequency of the radio wave (called the "carrier") before the message is sent on it.

Noise is an important concept in communication theory. Noise is defined in communication theory as any signal that interferes with the message being sent and is an undesired disturbance in a communication system. Radio static is a form of noise. Dirt on a camera lens is noise also. In a double-exposed picture, the unwanted scene, no matter how beautiful it may be by itself, is noise as far as it applies to the communication.



Any signal that interferes with the message being sent is noise

chapter 2

communication systems

Communication systems can be subdivided in many ways. One subdivision might be a real-time system in which the receiver is getting the message in the shortest possible time. One person listening to another on a telephone would be hearing his message in real time. If the message was sent via a tape recording, it would not be a real-time message. Books and motion pictures are also messages that are not sent in real time.

The majority of real time communication systems are electrical or electronic in nature. Telephones, telegraphs, radio, and television are examples of real-time systems (except for delayed transmission of recordings).

The earliest real time electrical communication system was the telegraph. It enabled messages to be sent over distances in minutes that would have taken days or weeks of travel otherwise.

A simple telegraph system is shown in Figure 1. This telegraph is a simplex arrangement and is useful only for transmitting information from a remote location. Since it allows no communication to the remote location, no clarification of a message is possible.

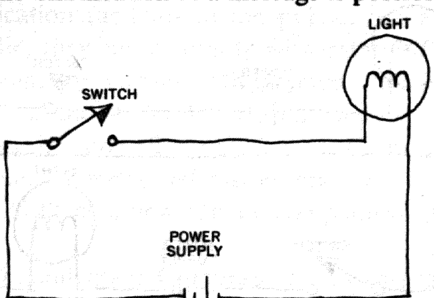


Figure 1 Simplex Telegraph

An example of another type of telegraph system, in which transmission and reception can be effected at either end, is shown in Figure 2. In this system, which is known as "half duplex," closing either signal key will cause both receivers to react. Because both receivers are affected by closing a key, it is possible to send messages only one way at a time.

If the sending and the receiving stations must send messages to each other simultaneously, a full duplex telegraph system must be used. An example of such a system is shown in Figure 3.

Electrical communication systems have one thing in common: they require connection through electrical conductors. This is true of telegraphs, telephones, and public address systems.

In electrical and electronic systems, there are two basic modes of communication — analog and digital.

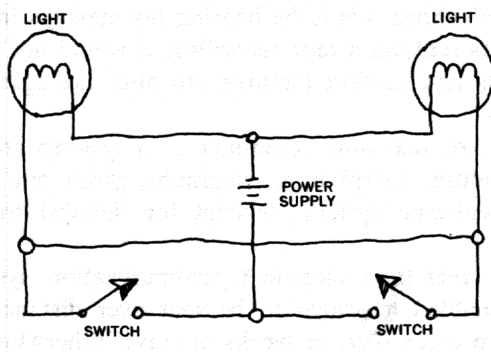


Figure 2 Half Duplex Telegraph

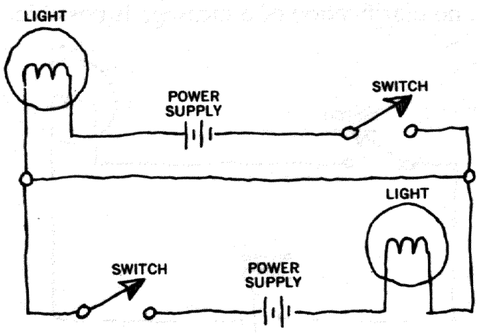


Figure 3 Full Duplex Telegraph

As in other electrical and electronic systems, an analog signal is the proportional electrical equivalent of the original communication. In a simple telephone circuit for example, the current is proportional to the intensity of the sound waves striking the mouthpiece. Public address and similar sound systems are analog communication systems.

Digital communication systems do not use a proportional electrical relationship. Rather, they are based on limiting the system to the presence or absence of electrical energy. The simplex telegraph shown in Figure 1 is an example of a digital system. The circuit switch can either be closed so that current is present and the signal device is actuated (this condition is known as the "one state"); or the switch can be open, so that current is not present and the signal device is inoperative (known as the "zero state").

Another way of differentiating communication systems is by the medium used to convey the signal. Signals can be sent along wires, radiated through space, sent along light beams, pumped through waveguides, or by a combination of these means.

Communication systems in the United States usually use wire or radio broadcast. Small private systems such as intercoms and private telephone systems generally use wires. Large commercial systems such as radio and television stations use broadcast links, not only to the home, but frequently for coverage of a remote event where it would be impractical to run electrical cables.

In many circumstances, when a message must be transferred between two or more terminals and the distance and the environment involved make it impractical for a user to supply communication facilities himself, he solicits the services of a communication common carrier.

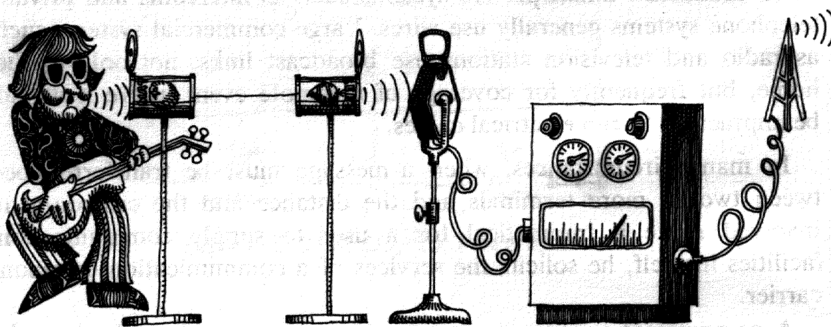
A communication common carrier is a company whose business is to supply communication facilities to the public. Since common carriers serve the public, they must comply with established regulations. A communication common carrier with interstate facilities comes under the jurisdiction of the Federal Communications Commission as well as the state organizations all are subject to. Common carriers can carry voice, facsimile, telemetry, television, and data messages. There are approximately 2800 common carrier companies operating in the United States.

The Federal Communication Commission was established in 1934 to regulate interstate and foreign communications by wire and radio. Its authority includes amateur radio, commercial radio, common carriers, television, and citizens' band radio. It regulates operators' licenses, classifies radio stations and delineates their services. By its actions, it insures that the services offered will be in the best interest of the public.

The FCC regulates common carriers by requiring that at least 30 days prior to availability, each company submit a schedule outlining those services the company intends to offer.

The two basic means most frequently used by common carriers for transmitting information are over wire conductors and by broadcasting radio waves. In a typical common carrier hookup, the transmitter and the receiver may be linked by wire-carrier links that in turn are connected by a radio link. From the standpoint of getting the information to the receiver, the composition of the communication hookup is not important, except to the extent to which it adds noise to the transmitted signal. Methods to overcome noise will be discussed in detail later.

Before discussing some of the larger common carrier communication companies, it is important to define and explain some of the terms used in this book.



In getting information to the receiver, the composition of the hookup is not important

Channel Grades — If we are going to transmit a desired message over a channel, the channel must be able to carry enough information to allow the receiver to understand what is being transmitted. If, for example, we wish to send music over a communication system, any owner of high fidelity music equipment will tell us that in order to hear everything that the orchestra is playing, we need a greater "frequency response" than we would find in a telephone earphone. This is an audio term describing the ability of equipment to handle signals within a determined range of frequencies. A typical music system might have the ability to handle frequencies from 20 Hertz to 20 Kilohertz. In communication terms, we would call this the *frequency range* of the equipment.

In a music channel, you need considerably more information to get a total message to the receiver than you do to get a spoken message. Unlike an orchestral piece, in which you must listen to the different instruments interacting with each other, transmitting a voice message means only that the words have to be understood on the receiving end of the communication link. The words can sound "tinny", for instance, without necessarily losing intelligibility. Since a normal human voice does not have the total tonal capabilities of an orchestra, transmitting a voice message requires a smaller frequency range than does transmitting an orchestral selection. Communication engineers have categorized this difference by establishing channel grades.

A music-grade channel is one suitable for the transmission of the total sounds of an orchestra, as heard by the human ear. A voice-grade channel is suitable for the transmission of speech, usually having a frequency range of from 300 to 3,000 Hertz. Channel grades are often classified in communications as either voice grade or wideband. In this terminology, any communication channel with a frequency range greater than that of a voice grade channel is termed a wideband channel.

Channel Capacity — In order to transmit information in its correct grade (e.g. voice, music, television etc.), we must determine whether we can establish a channel with sufficient range to allow the necessary information to be transmitted without being wasteful. This is particularly true in radio broadcasts, which require the use of upper and lower frequency limits within the radio spectrum. In such circumstances, the channel's frequency range is determined in part by the frequency of the carrier. It has been found that the frequency of the carrier must be at least double that of the highest frequency in the message being set. Therefore, if you had a message with a frequency range of 60 Hz to 6 MHz, you wouldn't broadcast it on a 4 MHz carrier; conversely, you know that with a 4 MHz carrier, the maximum frequency you can transmit is no more than 2 MHz. Thus, by knowing the band width of your channel, you can determine the maximum and minimum frequencies you can transmit; that is, you can determine the amount of information you can send. This is known as the channel capacity.

For example, a channel with sufficient bandwidth to transmit music would have a channel capacity substantially in excess of the requirements for a voice-grade channel. To use such a bandwidth merely to transmit one person's voice only would be highly wasteful; however, there is room in such a channel bandwidth for several voice grade channels. Therefore, in some communication systems, the capacity of a specific channel or system under discussion is often designated in terms of how many voice-grade channels it can contain.

Carrier Modulation — We have previously defined a carrier as the radio wave of a certain frequency before a message is sent on it. In order to transmit information, we have to put our message on this carrier. The process by which this is done is known as "modulating" the carrier.

There are three basic forms of carrier modulation:

a. Amplitude Modulation (AM),
b. Frequency Modulation (FM),
c. Phase Modulation (PM).

Each method can be used to transmit any signal that falls within the capacity limitations of the channel it is being used for.

We are most familiar with amplitude modulation (AM). It is employed in the majority of standard broadcast radio stations in the United States and in the video portion of a commercial television signal. Amplitude modulation is accomplished by changing the maximum strength of the carrier, with time, to conform to the change of the signal being transmitted with time.

Figure 4 illustrates this principle in detail.

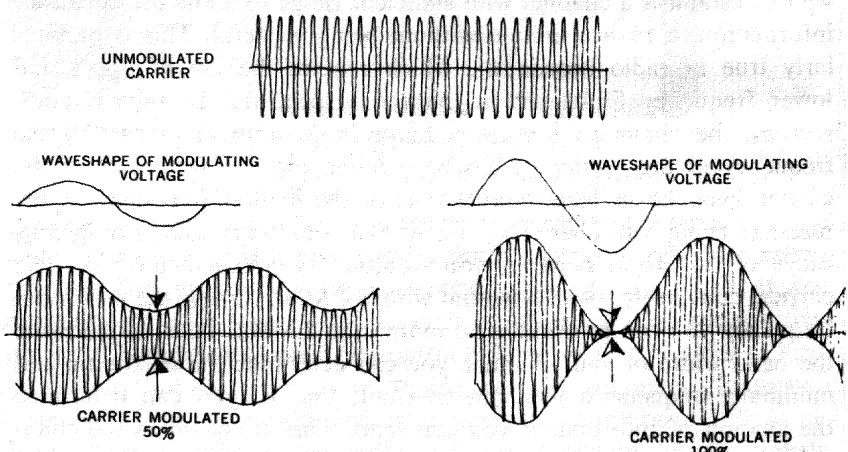


Figure 4 Amplitude Modulation

Frequency modulation is also used for commercial broadcasts, but unlike AM, the frequency modulated signal is broadcast with a constant amplitude. Instead of varying the strength of the carrier, its frequency is altered to conform to the modulating signal. This is shown in detail in Figure 5.

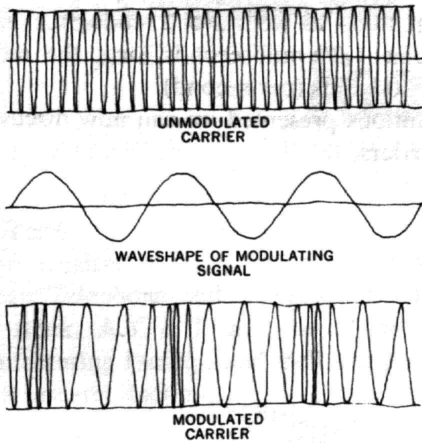


Figure 5 Frequency Modulation

Phase modulation is a less known form of carrier modulation. Instead of modulating the amplitude or frequency of the carrier, the phase of the carrier wave is modulated. Frequency and phase are not independent; therefore when the frequency of a carrier is varied, the phase is affected and when phase is varied the frequency is affected. In frequency modulation the frequency deviation is proportional only to the amplitude of the modulating signal; in phase modulation the frequency deviation of the carrier is proportional to both the amplitude and frequency of the modulating signal.

Pulse Coding — A signal in a communication need not be analogous to speech or music. The click of a telegraph key can be a signal. Anyone familiar with the Morse code knows that the letters of the alphabet, numbers, and some punctuation symbols are represented by a series of dots and dashes. Considering our simplex telegraph (Figure 1) we can see that the message can be represented as interrupted current flow. The message is thus made up of a series of electrical pulses. This is referred to as pulse coding, and will be discussed in greater detail in the section concerning data communication codes and formats.

Baud Rate — In data communication, a fixed amount of time is devoted to sending a pulse, known as a binary digit or "bit", (which can be either a positive pulse, as a telegraph dot, or a blank, as a telegraph pause). The number of information bits that can be transmitted in one second is the baud rate. By definition, a baud is the reciprocal of time in seconds occupied by the shortest element of the code being transmitted. If we have a code, for example, with its shortest signal element 20 milliseconds long, the modulation rate of the code would be 50 bauds (per second).

With these definitions presented we can now discuss some representative common carriers.

chapter 3

communication facilities

THE BELL SYSTEM

The Bell System consists of many companies of which the parent company (American Telephone and Telegraph Company) owns part of or all of the stock. A.T. & T. also owns the Western Electric Company, a manufacturing facility, and Bell Telephone Laboratories, an engineering and research facility.

The Bell System offers many communication services. Most independent companies also offer these services, although they are not directly affiliated with the Bell System.

The following is a brief outline of the services offered by the Bell System*.

Wide Area Telephone Service — (WATS)

This service permits a subscriber to place an unlimited or metered amount of outgoing calls within a specified radius at a flat rate. Monthly charges are based upon the radius of area covered and the amount of calls placed.

*Other communication companies offer services equivalent to the Bell System. We are describing the Bell System services as representative of those offered.

DATAPHONE® Services

Dataphone Services provides for the transmission of data using the regular dial-up telephone network. This is done by the use of modems (devices that modulate and demodulate signals) and the same dial telephone network used for local and long distance voice communication. The modem, which is connected electrically to the line, converts binary information available from digital equipment into tones, making it suitable for transmission over voice grade lines. In a like manner, the modem converts tones received on the voice grade line into binary data and makes them available to digital equipment.

The equipment that the Bell System supplies to accommodate these requirements is called a Data Set and varies in model depending upon the degree of automation and data required. Recent changes in regulations permit the connection of non Bell System Data Sets to the dial up network via a Data Access arrangement available from the Bell System.

Telpak

Telpak service is a pricing arrangement that permits a wide-band communication channel between two points. This wide band can be used as the subscriber wishes. It can be broken down into many voice grade lines or used as one high-speed data line. There are four basic Telpak services available.

Telpak A	equivalent to 12 voice grade lines
Telpak B	equivalent to 24 voice grade lines
Telpak C	equivalent to 60 voice grade lines
Telpak D	equivalent to 240 voice grade lines

Channel capacities similar to that of Telpak are also available with private microwave systems.

Private Line Voice Service

Private line voice service provides a private voice-grade line for the exclusive use of a particular subscriber. This service can be used for either voice communication or data in transmission using a modem. The charges for private line service are based on a per mile, per month basis.

Private Line Teleprinter Service

Private line teleprinter service provides circuits for the exclusive use of particular subscribers. Keyboard printers, paper tape punches, and paper tape readers are used with these services. One circuit can connect two or more teleprinter machines.

TELETYPE EXCHANGE SERVICES (TELEX & TWX)

Teletypewriter Exchange Service provides direct-dial-point-to-point connections using teleprinter equipment, such as keyboard printers, paper tape readers, and tape punches. Facilities are also available to permit computers to interface to these services. These are Western Union services.

GENERAL TELEPHONE AND ELECTRONICS

General Telephone and Electronics (G.T.&E.) is a large communication and manufacturing organization with facilities in many areas of the country. It offers most of the services available from the Bell System.

WESTERN UNION

The Western Union Telegraph Company furnishes communication services by wire and microwave radio throughout the United States. It also provides custom built private-wire systems, facsimile systems on a lease basis, Telex, and TWX. (TWX was purchased from Bell System in 1971.) Western Union also offers leased-line services for data transmission as well as teletypewriter service.

COMMUNICATION FACILITIES AVAILABLE

The facilities offered by the communication common carrier can be broken down into the three following groups:

Telegraph Channels

These are lines that are separate from the voice grade network and can accommodate speeds up to 75 baud. A dc current modulation technique is often used to transmit data. Most common carriers will not allow dc current modulation to exceed 75 baud on any of their facilities. If a higher data rate is desired on these lines, a modulator-demodulator (modem) that can bring the data up to 150 baud is required. 300 baud is usually the limit for this grade facility.

Voice Grade Channels

Voice grade channels are lines that are part of a voice grade telephone network and have all the advantages of that network. However, private voice-grade lines are available that are of higher baudwidth, if desired.

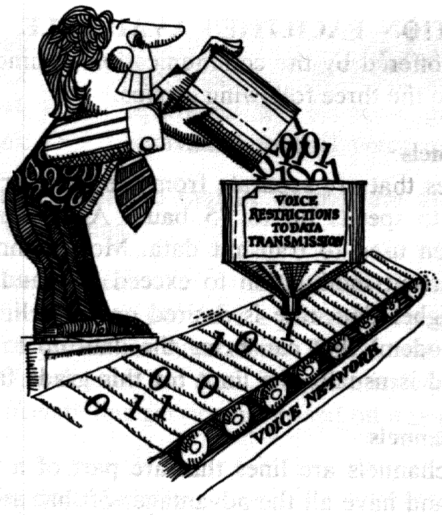
Wide Band Channels

Wide band channels are facilities that the common carriers provide for transferring data at speeds of from 19200 baud up to the 1 million baud region. The availability of these facilities depends upon what equipment the common carrier has in the subscriber's geographical area. There are many modems available for use with common carrier facilities, most of which are supplied by the common carriers. Tables of common carrier facilities and characteristics may be found in Appendix A.

SPECIALIZED COMMON CARRIERS

The current facilities that the common carriers offer are basically designed for voice communications. In spite of the carriers' great success in adopting these services to data transmission, the basic characteristics of these existing facilities prohibit them from being responsive to the increasingly diverse needs of the computer and terminal communications market place.

Recognizing these limitations and the increase in their effect, a number of specialized common carriers have been working on plans to offer networks specifically for data transmission. Their goals are to offer an improved billing structure and lower error rates.

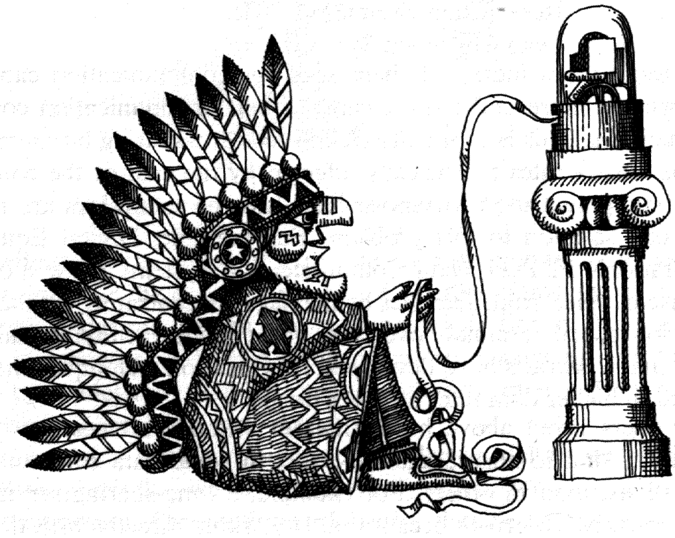


A general common carrier is not the most efficient method of transmitting data

chapter 4

data communication

A large subdivision of "communication" is the field of data communication. Unlike the total communication discipline, data communication is primarily concerned with the transferral of numerical or instrumentation messages. Thus, large manufacturing firms interested in inventory or payroll information, scientists in laboratories who want to use the facilities of a large computer located halfway across the city, and motel chains that are interested in making advance reservations for their customers at other cities, use data communication systems. The increasing influence of technology on modern civilizations requires greater use of data communication.



The increasing influence of data communication

Previously, we discussed pulse code messages; in data communication, pulses are used because they are easy message elements for electrical circuits to handle. There are several forms of pulsed messages, but the most basic utilizes the principle of binary numbers. Since binary numbers can be represented by an electrical signal (1) or a lack of electrical signal (0), it is evident that such a numbering system would be desirable in electrical circuitry; the presence or absence of electrical energy is all that would need to be detected.

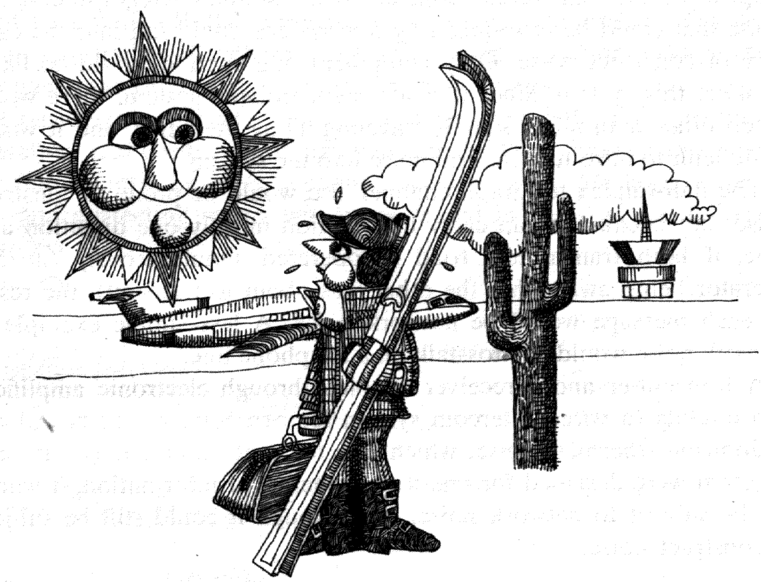
Digital data communication is especially useful because most electronic computers use digital counting. Thus, by using communication facilities that can handle digital data, electronic computers can be interconnected and used to help process and transmit data.

With the rapid advancement of computer development has come the capacity of computers to handle many users, either simultaneously or with such rapid response time that it is imperative that the computer system be capable of communicating with many terminals simultaneously. This requirement in turn has resulted in common carriers offering more reliable and more highly efficient data services. Typical users of computer based communication equipment are:

Inventory Control
Time-Sharing Service Utilities
Management Information Systems
Message Storage and Forward Systems
Reservation Systems
Data Collection Systems

As these users increased their need for communication capability, they became more and more aware of their communication cost on a per line basis. This is especially true in the time-sharing business where the customer's rates are directly affected by the cost of the communication equipment used to service him. Line concentrators are proving to be one solution to this problem. For example, Digital Equipment Corporation's PDP-11 Data Communication Systems allow more efficient use of voice grade facilities by message concentrating the data from many low-speed terminals over one voice grade facility. In addition, the common carriers have developed transmission techniques that have resulted in higher data rates at lower prices.

The users listed above have different criteria for message quality. It is more vital to ensure against errors in some data communication applications, than in others. For example, a time-sharing service user is less sensitive to errors because he is working directly with the computer and thus is an integral part of the data stream. He can catch errors as they occur. On the other hand, errors in reservation systems could be very serious. A traveler could be issued a ticket for the wrong airplane flight and find himself stranded or on a flight going to the wrong destination. All systems are subject to errors; some have a lower tolerance level than others. Finding the acceptable tolerance level is very important because the higher quality the message has to be, the more redundant the data communication system must become. If redundancy is added to equipment, the cost of the total system increases.



Errors in reservation systems can be serious

To understand the place of errors in data communication systems, we must consider the concept of *noise*. Communication theory defines noise as an undesired disturbance in a communication system, whereas a signal is defined as a desired disturbance. To this extent, any generation of an undesired signal is noise, including malfunctions of the sending and/or receiving hardware.

Noise produces message errors. In order to define our types of error more precisely, we will necessarily subdivide our noise into different categories. These will be arbitrarily classified as:

a. Construct Noise — noise caused by equipment malfunction

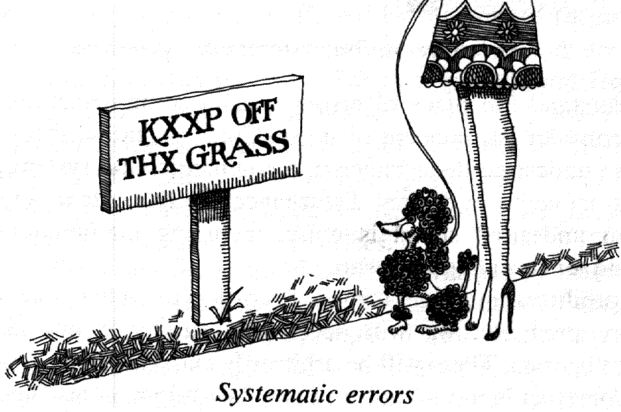
b. Natural Noise — noise caused by natural phenomena such as thermal emission, static, etc.
c. Network Noise — noise caused by interaction of another signal with the signal under consideration

Errors caused by any of the above types of noise can be either random or systematic. The probabilities of error introduced by each of the above noise types however, will not be equal. The probability of a particular type of noise present will depend upon the type of equipment being used.

Let us examine representative communication systems. Perhaps the simplest device that we can consider is the simplex telegraph channel. Noise that could be introduced by this system would fall into the category of construct noise. Only equipment malfunction would be likely to affect this system. Since it is a closed, isolated system, there would be no other man-made signals traveling along the wires, and it would be difficult to introduce natural noise into the system.

The half-duplex telegraph channel also would be prone to construct noise; but because it can carry information in only one direction at a time, if both transmitters tried to operated simultaneously (neither operator being aware that the other was about to transmit), the result on each message would be network noise. A less simple example of network noise would be crosstalk on a telephone line.

A transmitter and a receiver coupled through electronic amplification circuits (a wired intercom system for example) would be subject to Johnson (thermal) noise, which is a form of natural noise. If such a system were designed for one-way exchange of information, it would not be subject to network noise, even though it could still be subject to construct noise.



Systematic errors

A transmitter and receiver coupled through a radio link would be subject to all three classifications of noise. Noise may also be classified as random, periodic, or systematic. Systematic noise, a special case of construct noise, is a form of noise inherent in imperfect equipment. For example, if the signal for the character "e" were always replaced by the signal for character "z" in a message because of equipment failure, this would be a systematic error. Interestingly enough, such systematic errors, once determined, do not interfere totally with a message.

Random noise is characterized necessarily, by a lack of organization. Precisely because it is random, its periodicity and amplitude are hard to predict. If the noise remains below a determined threshold level, it usually will not interface with a signal, if it does not, system parameters must be adjusted to minimize its effects.

Considerable thought and effort have gone into the prevention, detection, and correction of errors in signals. These have formed one of the fundamental concepts in communication theory. Methods of detecting and correcting errors will be discussed in detail in the following chapters.

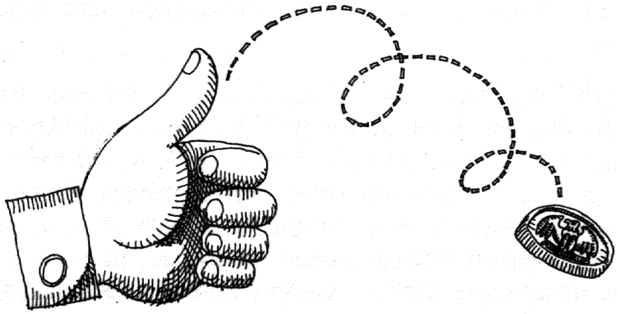
chapter 5

data communication techniques and systems

There are several techniques used for the transfer of data communication signals. Each has its particular advantages, and the trade-offs of each type will be discussed in turn.

Standard data communication messages are sent in some form of pulse code. Data communication codes are discussed briefly at this time but will be covered in detail in Chapter 6.

As discussed previously, there are several varieties of pulsed codes that are used in the transferral of data in digital form. Binary signals, by their very nature, are natural elements for digital data codes. Such codes are said to be in "binary format".



Binary code decision making device

A binary formatted code can represent different symbols only by allowing sufficient binary elements for each symbol. If we think of one binary digit (or "bit") representing each symbol, we have only two choices: one symbol represented by the "on" state, the other represented by the "off" state. With such an arrangement, we could let the "on" or one-state represent "no" and the "off" or zero-state represent "yes". While it would be difficult with such an arrangement, we could convey messages of a very limited nature from a remote station (such as the answer to "Is the temperature at your station over 70°F?").

If, instead of using one binary digit for our character, we use two, we now have more characters to choose from. Our choice for a one-bit code was limited to two: 0 or 1. Our choice for a two-bit code is four: 00, 01, 10, or 11. If we choose a three-bit code, our choice is eight: 000, 001, 010, 011, 100, 101, 110, and 111. It can be shown that for a code with a character makeup of n bits, the number of characters available will be 2^n. In communications parlance, instead of calling these codes one-bit codes, two-bit codes, etc., they are called one-level codes, two-level codes, etc.

Although we could assign any arbitrary meaning to a code character, it is more practical for the majority of operations to let the characters represent letters of the alphabet, numbers, punctuation marks, and spaces. In addition to these, some special codes use characters for other meanings. Codes and formats will be discussed in detail in the next section of this book.

In order to transmit code characters, it is necessary to arrange their elements in a way that will allow their reception without uncertainty. There are several techniques by which this may be done; these fall into two broad categories: serial data transmission and parallel data transmission.

In parallel data transmission, each element of a code has its own channel so that the total character is transmitted at the same time. This means that a five-level code, for example, would have five channels in a parallel transmission setup. For an electrical impulse transmission, where each channel would be a pair of conductors or a conductor and return through a common ground, such an arrangement could use many more conductors than would be practical for a total communication hookup. A five-level parallel data transmission arrangement is represented in Figure 6.

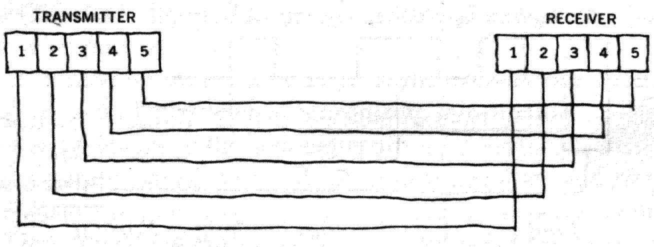


Figure 6 Parallel Data Transmission

In serial data transmission, each element of a code character is sent in turn, rather than sending all elements at once. This technique allows binary coded multi-level data to be transmitted over a single channel. In electrical impulse transmission, this means that the total communication hookup can be a pair of conductors or a single conductor and ground. A five-level serial data transmission arrangement is represented in Figure 7.

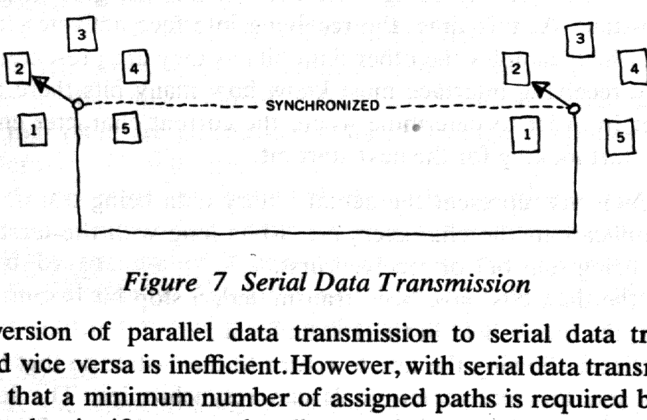


Figure 7 Serial Data Transmission

Conversion of parallel data transmission to serial data transmission and vice versa is inefficient. However, with serial data transmission, the fact that a minimum number of assigned paths is required becomes increasingly significant as the distance between the transmitter and the receiver increases. There are two basic forms of serial data transmission: synchronous and asynchronous. Each will be discussed in detail.

ASYNCHRONOUS SERIAL TRANSMISSION

Asynchronous serial data transmission is the technique used by most electromechanical serial devices such as teletypewriters. With this technique, each character consists of three parts: a start bit, the data bits, and a stop bit. An example of this format is shown in Figure 8.

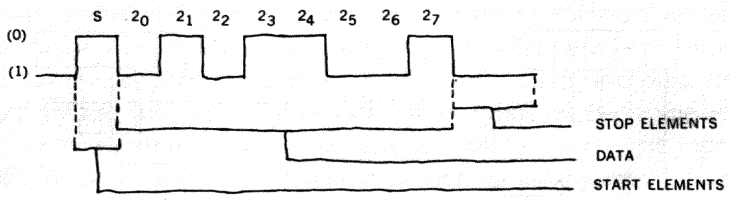


Figure 8 Asynchronous Technique Format

A start bit is a line state (usually a zero) that lasts for 1-bit time and is used to indicate the beginning of a character. The beginning of a character can be detected through adherence to two basic rules.

a. When data is not being sent on the line, the line is kept in the 1-state.
b. After the last data bit is transmitted, the line will return to the 1-state for at least 1-bit time.

Using these two rules, the interface can detect the start bit when a character is not actively being received, and the line goes from the 1- to the 0-state. At this time, the receiving interface activates (enables) a clock, which samples the other data bits as they are presented on the line. The receiving interface must know how many bits there are to a character in order to determine where the current character ends, and when to start looking for the next start bit.

The data bits represent the actual binary data being transferred. In many applications the characters are 8-bits long with the least significant bit being sent out or received first.

After the data bits have been transmitted, a stop bit is sent. During this time, the line is held in the 1-state for a period that lasts 1-, 1.42-, or 2-bit times allowing the line to return to the 1-state so that both the transmit and the receive terminals can resynchronize. The length of time that the line stops in this state (i.e., 1-, 1.42-, or 2-bit times) depends on the amount of time required for the equipment to resynchronize. Most electromechanical equipment uses the 1.42-, or 2-unit stop codes; however, newer mechanical equipment, such as Teletype Model 37 uses only 1-stop bit. The line will remain in the 1-state until another character is transmitted. This stop time insures that a new character will not follow for at least 1-, 1.42- or 2-bit times.

The asynchronous serial data transmission technique has the following advantages

a. Can easily be generated by electromechanical equipment. (e.g., Teletype keyboard)

b. Can easily be used to drive mechanical equipment. (e.g., Teletype printer)

c. Characters can be sent at an asynchronous rate because each character has its own synchronizing information.

The disadvantages of this technique are

a. Separate timing required for both transmitter and receiver.

b. Distortion sensitive because the receiver depends upon incoming signal sequences to become synchronized. Any distortion in these sequences will affect the reliability with which the character is assembled.

c. Speed limited because a reasonable amount of margin must be built in to accommodate distortion.

d. Inefficient because at least 10-bit times are required to send 8-data bits. If a two-unit stop code is used, it takes 11 bits of time to transfer 8-bits of data. Of course, more than 8-data bits can be transferred to improve efficiency, but this is usually not the case.

SYNCHRONOUS SERIAL TRANSMISSSION

In the synchronous serial technique, a serial-bit stream is sent over the line in the same manner as the asynchronous serial bit stream except that there are no start-stop bits with which to synchronize each character.

In this technique, the entire block of data is synchronized with a unique code which, when recognized causes the receiver to lock in and set, and by the use of a counter, to count the incoming bits and assemble a character. As in the asynchronous technique, the receiver must know the number of bits to a character.

Timing Signals

Unlike the asynchronous technique, a synchronizing signal must be provided along with the data-bit stream. This signal can be generated by the transmitter or by a separate source that the transmitter uses for transmit timing. In either case the data must be transmitted and received synchronously with a common clock. The synchronous technique format is shown in Figure 9.

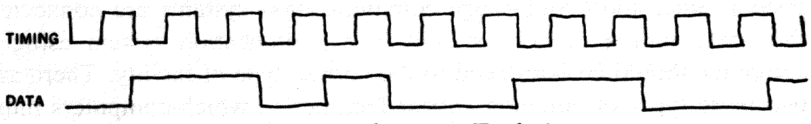


Figure 9 Synchronous Technique

In the format shown, the transmitter presents data to the line on the negative transition of the timing signal and the receiver samples the data line on the positive transition.

The advantages of the synchronous serial techniques are

a. Common timing source can be used for both transmitter and receiver.
b. Receiver does not require clock-synchronizing logic as the asynchronous technique does.
c. Highly efficient because there are no bit times wasted with the use of start and stop bits. All bits on the line are data, with the exception of the synchronizing pattern at the beginning of the bit stream.
d. Low distortion sensitivity because the timing is provided along with the data.
e. Higher speeds are achievable because of the low distortion sensitivity.

Disadvantages of the technique are

a. Characters must be sent synchronously, not asynchronously as they become available (which is desirable for most real time and mechanical applications).
b. One bit time added to or missing from the data-bit stream can cause the entire message to be faulty.
c. The common-carrier equipment to accommodate this mode of operation is more expensive than the equipment required for asynchronous modes of operation.
d. Mechanical equipment cannot transmit or receive this format directly.

Although it is possible to use a synchronous communication facility for an asynchronous format, with start-stop bits, there is limited application for such an approach.

As mentioned in Chapter 4, electronic computers are often connected into a communication system to help transmit and process digital data. By using computer systems to concentrate data from many low-speed terminals over one voice grade facility, significant improvements can be made in the efficiency of a data communication system. Since most long-range communication systems are connected through common carrier facilities, a communication system using a computer should be interfaced to the correct type of facility. There are two basic types of common carrier facilities to which computers must be interfaced: asynchronous serial and synchronous serial. We have already pointed out the advantages and disadvantages of these two types of facilities. Based on these advantages and disadvantages, Table 1 shows typical speeds and applications of these two techniques.

As shown in Table 1, there are three basic communication applications to be solved by the computer communications engineer.

Low speed terminal equipment, such as teletypes.

Medium speed terminal equipment.

Intercomputer communications.

TABLE 1

SPEED	ASYNCHRONOUS	SYNCHRONOUS
Low 0 to 300 Baud	Electromechanical terminals such as keyboard printers and teletypes.	Operations tend to be asynchronous at these speeds.
Medium 300 to 5,000 Baud	Unbuffered terminals such as paper tape readers and punches, card readers and line printers.	Buffered terminals such as displays, and buffered card readers, and line printer configurations.
High 5,000 Baud and up	Not frequently used.	Intercomputer communications.

The application and speed of the lines are not the only parameters to be considered in the design of line interfacing hardware. The number of lines to be serviced is equally important because significant per-line cost savings can be realized when interfacing many lines. Depending upon the number of lines, common character memory and synchronizing techniques can be implemented in the design. Some of these techniques are expensive initially, but when used for a number of lines, the price on a per-line basis is low.

Asynchronous Line Interfaces

Interfacing to an asynchronous serial line is relatively straightforward for a logic design engineer. As an example, let us use the standard Teletype Baudot Code Format shown in Figure 10. The Baudot Code will be explained in detail in the following Chapter.

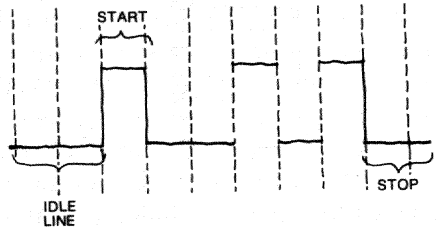


Figure 10 Example of Code Format

The format above shows that the beginning of a character is clearly defined by the leading edge of the start bit. Using this leading edge as a reference, the interface must synchronize so that it samples each bit in the middle.

Once the beginning of a character has been detected, the end of a character is easily determined by counting the number of bits sampled or by detecting the start bit shifted out of the last position. To detect the beginning of a character, the interface needs a flip-flop circuit that holds the active/inactive status. This flip-flop gets set when the start bit of a character is detected and is turned off when a character has been received. If the interface is inactive, and the data line goes from the one-state to the zero-state, the start bit of the incoming character is indicated.

There are two methods that can be used to sample the middle of each bit:

a. Gateable clock, which has a start-up time equal to ½ a bit time, and whose frequency equals the baud rate of the line being sampled.

b. Counter, which causes a sample to be taken at overflow.

The gateable clock approach is seldom used because of the inherent variation in such a low speed clock, which would cause it to drift from the center of each bit. The counter, on the other hand, will not necessarily sample the middle of the bits, but if driven by a crystal clock, which is divided by a counter, will always sample within a certain percentage of the center of each bit.

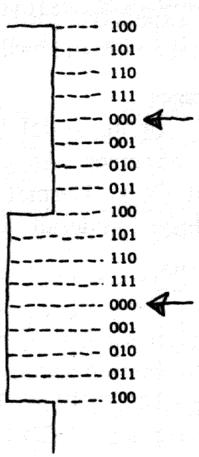


Figure 11 Common Sampling Design

A common sampling design uses a 3-bit counter (count of 8) that gets set to 4 (100) when the leading edge of a start bit is detected. The counter is incremented at 8 times the bit rate using the overflow to generate a sample. This approach is outlined in Figure 11.

Using the preceding discussion as a reference, we can draw the following block diagram and flow chart to define a typical asynchronous line interface (See Figures 12 and 13).

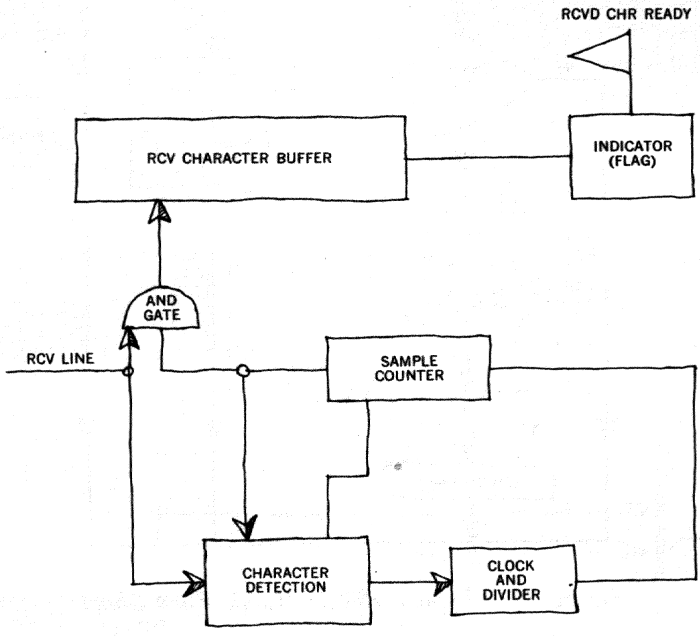


Figure 12 Asynchronous Line Interface

When designing an interface for handling a few lines (typically less than 6), a duplication of the complete interface logic shown in Figure 12 is sufficient. The cost and space saving is seldom significant enough to justify eliminating redundancy. This is because the computer interfacing is a high percentage of the total cost, and the saving of a few discrete components is probably not significant. However, if all lines are the same speed, the clock can be made common to all interfaces.

As the number of lines increase from 10 to 100 and greater, it becomes obvious that a significant cost saving can be realized by eliminating as much redundancy as possible.

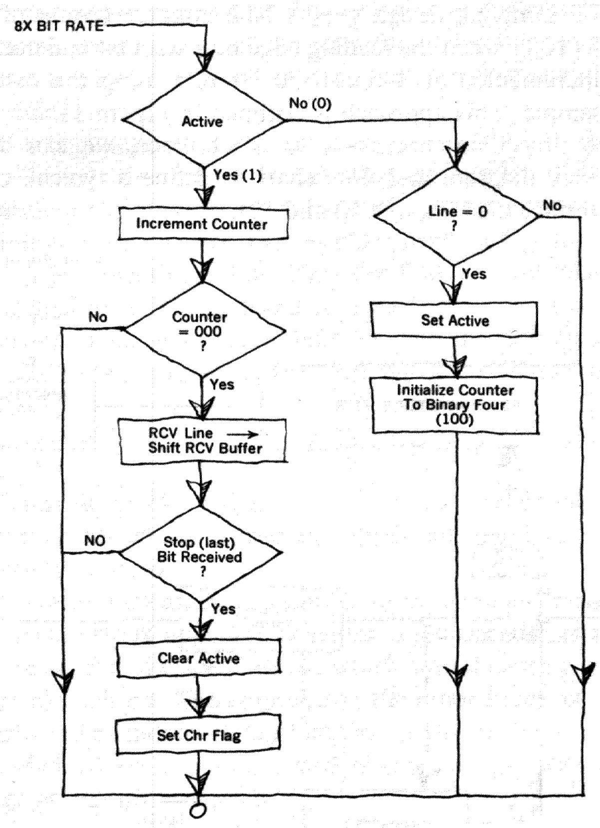


Figure 13   Typical Receive Logic Flow Chart

COMMON MEMORY SYSTEMS

One of the first steps to be taken to reduce the cost of a serial line interface is to eliminate the receive character buffer and assemble characters in a memory that is common to all line interfaces. Two such common memories are: 1) delay lines and 2) core memories. A common delay line approach is economical for many lines, but the economy decreases sharply when the same interface is used to handle fewer lines than its designed capacity.

Core memory is more expensive than delay lines but its higher speed and random access capability make it a useful common memory when designing a system to handle many medium or high speed lines. Core memory also permits the random mixing of line speeds because location can be accessed at different rates.

When a common memory design approach is used, the character buffer can be stored in the common memory, thus eliminating hardware registers. However, start bit detection, sample synchronization, and character assembly detection must remain as hardware unique to each line. A typical multi-line common memory system is shown in Figure 14. Figure 15 shows a typical line unit.

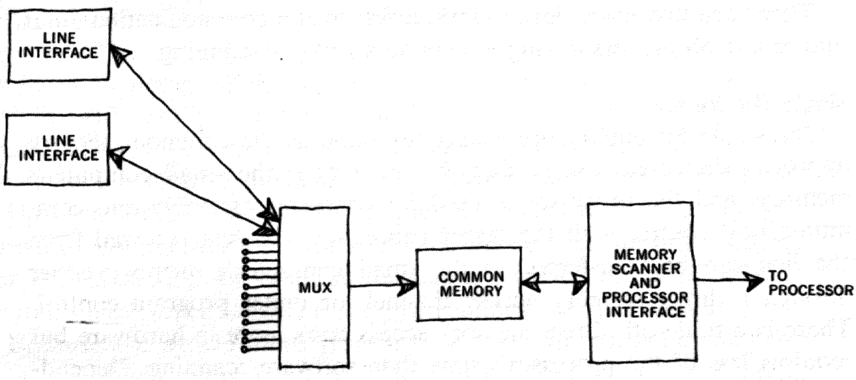


Figure 14 Typical Multi-Line Interface

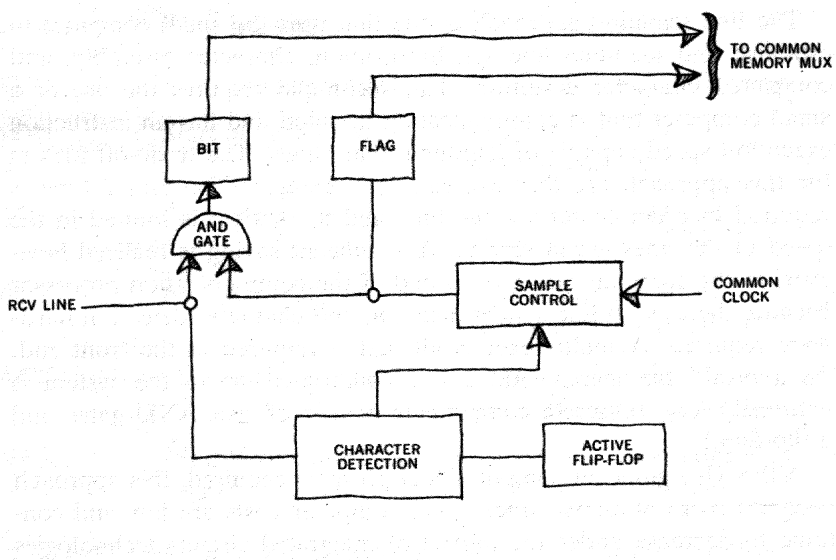


Figure 15 Line Interface in a Common Memory System

Small Processors

Another technique being used to handle many asynchronous lines is the small computer that serves as a line controller for the large processor. The small computer not only can serve as a line scanner and controller, but can handle a number of "supervisory" tasks that would normally be done by the large processor, such as error detection, character echoing (on a full duplex line), user validity checking, etc.

There are two basic design approaches to the communication front end of a small processor: single bit buffers and line scanning.

Single Bit Buffer

The single bit buffer approach is the same as the common memory approach discussed, except that the memory is the small computer's memory, and the processor is used for character assembly and communication control with the larger processor. The bits received from the line can be transferred to the small computer's memory either through a direct memory access channel, or under program control. There is a trade-off: direct memory access costs more in hardware but requires less of the processor's time than software scanning. Depending upon how the small computer is used, a direct memory interface costs 10% to 50% as much as the small computer itself.

Line Scanning

The line scanning approach is one that uses the small computer to establish and maintain line synchronization, character assembly, and completed character detection. This technique requires the use of a small computer that is communication oriented and has an instruction execution speed capable of handling many lines. The trade-off factors for this approach are that a greater percentage of processor time is required in order to service the line, and the system is limited in the speed of the lines it can service. A significant saving is realized however, in the communication front end of the communication processor because there is no line synchronization and character detection hardware required. A multiplexer is all that is required at the front end. As a result, the incremental cost of adding a line to the system is extremely low. (Discrete components consist of two AND gates and a flip-flop.)

Although significant computational power is required, this approach becomes more attractive since small computer costs are low and continue to decrease under the impact of integrated circuits technologies and improved mass production techniques.

Synchronous Line Interface

Synchronous modulator-demodulators (modems) have permitted a higher rate of data transmission than asynchronous modems over a voice grade facility. The nature of these transmission techniques has also resulted in higher efficiency by eliminating the need for synchronizing information with every character.

The logic design of interfaces to a synchronous modem is considerably easier than the design of an asynchronous interface because there is no need for bit synchronization and sampling hardware. Most synchronous modems supply all the timing necessary to receive each bit as it is made available from the modem. The difficulty in designing a synchronous modem interface is to design in the capability of communicating in the message formats used in synchronous communications.

It is not the purpose of this chapter to discuss the format for synchronous communication, which will be detailed in the next chapter, but a brief description of these formats is outlined below in order to establish a reference for discussing synchronous interface design.

Because the synchronous transmission technique provides only bit recovery timing, there must be a way to establish character framing and message framing. This is accomplished by using codes (usually ASCII) that are assigned for synchronous message formating purposes. Listed on the next page are representative message codes.

Character	Meaning	Function
SYN	Synchronizing signal	Establish character framing
SOH	Start of heading signal	Precedes block message heading characters
STX	Start of text signal	Precedes block of text characters
ETX	End of text signal	Terminates a block of characters started with STX
ACK	Acknowledge signal	*Affirmative acknowledgement of a message received
NAK	Negative acknowledge signal	*Negative acknowledgement of message received

*ACK is sent by the station that received a message to the station that originated the message.

A typical message that might be sent between two devices (a terminal and a processor) is shown below.

Terminal	**Processor**
SYN	SYN
SYN	SYN
SYN	SYN
SOH	ACK
User Terminal N 4	SYN
STX	SYN
Req. Balance of Account No. 14325	SYN
ETX	SOH
LRC (check character)	To terminal No. 4
Idle Line	STX
SYN	Balance is $100
SYN	ETX
SYN	LRC (check character)
ACK	Idle Line
Idle Line	

A look at this message reveals that there are a number of characters used only for message formating. The only data that the processor is concerned with is "Request Balance of Account No. 14325" and that it came from Terminal No. 4.

If the processor has to do all character decoding of each message, filter out and check all control characters received, a significant percentage of processor time is spent doing communication formating when it could be better spent for data processing. If, on the other hand, the processor has an instruction set that can efficiently perform this function, it is justifiable for handling a few lines, where the price of such a controller is high on a per-line basis.

An alternate approach to relieve the processor of this overhead is to design this capability into the hardware, however, the expense involved in doing this for a few lines is extensive (i.e., it can double the cost of interface) and can result in limiting the application for which the synchronous interface can be used.

To design in special character recognition means the addition of a number of states and to design the interface further to permit redefinition of special characters (to be compatible with other formats) requires additional gating or even registers.

Once hardware is designed to be format sensitive, the market for that interface to applications and terminals that conform to the particular format is limited. To design the interface further in order to have flexible hardware format capability causes the price of the interface to become prohibitive to those users who operate with fixed formats.

A typical problem: If you design an interface that handles and strips all format and control characters, passing only headers and text onto the processor, how do you use this same interface to serve as an interface between two processors or a processor and a graphic terminal that requires uncoded binary data be transferred between processors? There are solutions to this (discussed in the last chapter) but they result in a cost factor that substantially increases the price of the interface.

One significant advantage in using a small computer as a multi-synchronous line controller is that it also relieves the large processor of message checking and acknowledgement and can communicate with the larger processor on a message block basis rather than a character at a time. It might even have direct access to the large processor's data file.

Speed of Synchronous Lines

The increasing availability of high-speed synchronous lines, (in the range of 50,000 baud) poses a new problem to computer manufacturers. As these services become more readily available and the costs (which are very high now) come down, there is going to be a need for a better solution than using the medium speed interfacing equipment to handle these high data rates. At the present time, the only solutions to handling these high-speed lines are:

a. Completely automatic interface for each line.
b. Small computer that services a few lines using less expensive synchronous line interfaces.

In either case, the cost of interfacing to a high-speed synchronous channel is high. One step that can be taken to keep the high-speed line interfaces to a minimum is to develop a standard message format for high-speed synchronous communications that all computer manufacturers can conform to, eliminating the need for costly hardware flexibility. The new United States ASCII code, if adopted by the industry, could be the beginning of such a standard.

There are certain considerations that must be made in designing a synchronous line interface. Typically, the problems listed below will form the core of the design parameters that should be considered. Quite often, the lack of resolving these considerations leads to such concessions, as having to use a full duplex facility for half duplex operation.

Problem: The synchronous nature of the interface requires you to guard against output characters being supplied to the interface belatedly.

Solution: Double buffer and detect late characters if the interface is serviced by software on a character basis. In any other case, the interface should flag (a character late) condition in order to prevent the processor from transmitting the remainder of a bad message. (The processor can terminate the bad message and retransmit.)

Problem: Because the received data are dependent upon the received carrier, the interface must detect an interruption in the received carrier in the middle of a message. These types of errors are covered by parity and redundancy checks, but if one bit is picked up or lost, the EOT (End of Transmission) will not be detected or the bad character framing might result in a false EOT code.

Solution: By monitoring the receive timing (or carrier), the hardware can detect a pause or fade in the modem carrier and flag the computer that the message is probably bad.

Problem: How do you guard against fake sync characters?

Solution: Most message formats have three sync characters; the message should be checked for at least two consecutive sync characters before considering the message valid.

Problem: If the sync characters at the leading edge were missed and as a result no message was detected, how does the computer know that a message was sent?

Solution: The transition of the modem carrier lead can serve as a possible indication that a transmission from the other end was not received. However, this is not a sound approach because of possible noise on the line, along with the unique real time considerations required.

The only end that is certain of an attempted transmission is the transmitter. Based on this, the best solution to this problem is to have the transmitting station transmit the message if it has not received an acknowledgement (ACK or NAK) from the receiving end within a fixed period of time. (Timeout)

chapter 6

digital communication codes and formats

As outlined in the previous sections of this text, the synchronous and asynchronous transmission techniques each pose a different message format problem due to the errors that occur through transmission.

Primarily, errors that occur during asynchronous transmission affect only those data that were being transferred at the time the environmental condition that caused the error was introducing errors into the system. This is because synchronization is established for each character. When synchronous communication techniques are used, it is possible for the entire remainder of a message to be bad once an error has been introduced into the system because the gain or addition of a bit stream will put the entire message out of framing.

Based on this fact, it is easy to conclude that when asynchronous techniques are used, corrections can be made by requesting transmission of specific characters in the message that had errors, parity errors, etc. However, this technique is not practical because the cost of implementing the unique character retransmission capability is prohibitive. The cost of an asynchronous system capable of transmitting unique characters would probably exceed a synchronous system with a speed so much greater, that it could afford to retransmit entire messages and still offer an effectively higher data rate over a given common carrier line.

The nature of asynchronous equipment, and the possibility of an entire synchronous message being bad, makes it common practice to correct a defective block by retransmitting the entire block.

In Chapter 4 we discussed the problem of noise which is the basic cause of errors, and it was through the introduction of noise as a parameter in the analysis of communication systems that communication theory came into being.

A total consideration of the theorems and mathematics of communication theory is beyond the scope of this text; however, it is possible to show a method in which we are using noise-induced errors as design parameters.

Rather than trying to take the extreme case of a communications channel with random noise, let us consider a simpler case, in which errors are caused by bursts of noise with a lower limit of repetition.

The noise that causes these errors occurs as impulses with a clearly defined minimum. During this minimum time, data can be transferred error-free. Figure 16, a typical communication channel wave, shows our hypothetical noise bursts (errors) that occur, for instance, at a worst case of 10 second intervals. In reality, the actual characteristics for this type of interface will vary with areas of the country, type of facility and transmission technique. Using this example, however, it becomes obvious that any block of data transmitted that is 10 seconds in duration, will have errors, therefore, it is important that each message block not be of extensive length. Typical block lengths are usually not in excess of 1,000 8 bit characters.

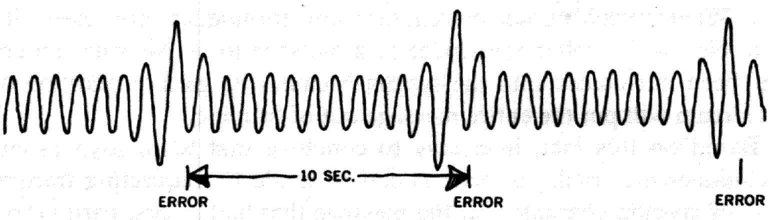


Figure 16 Line Errors

There are many codes used in communications today, and many new codes continually being developed for various applications. This section outlines some of the more popular codes being used in the communication industry, describes features that are unique to each and how they apply to terminal-to-computer and computer-to-terminal communications.

BAUDOT CODE

The Baudot Code is a 5-level code used only for telegraphs, keyboards, printers, punches and readers.

Although 5 bits can accommodate only 32 unique codes, two of the codes are figures (FIGS) and letters (LTRS). Prefixing the FIGS or LTRS code before other bit combinations permits dual definition of the remaining codes. This means that when a baudot terminal is interfaced to a computer, the software must maintain proper FIGS — LTRS status in order to interpret the necessary data properly.

CHARACTER		IMPULSE POSITION				
LOWER CASE	UPPER CASE	1	2	3	4	5
A	-	●	●			
B	?	●			●	●
C	:		●	●	●	
D	$	●			●	
E	3	●				
F	!	●		●	●	
G	8		●		●	●
H	#			●		●
I	8		●	●		
J	'	●	●		●	
K	(	●	●	●	●	
L	)		●			●
M	.			●	●	●
N	,			●	●	
O	9				●	●
P	0		●	●		●
Q	1	●	●	●		●
R	4		●		●	
S	BELL	●		●		
T	5					●
U	7	●	●	●		
V	;		●	●	●	●
W	2	●	●			●
X	/	●		●	●	●
Y	6	●		●		●
Z	"	●				●
LETTERS (SHIFT TO LOWER CASE)		●	●	●	●	●
FIGURES (SHIFT TO UPPER CASE)		●	●		●	●
SPACE				●		
CARRIAGE RETURN					●	
LINE FEED			●			
BLANK						

PRESENCE OF ● INDICATES MARKING IMPULSE
ABSENCE OF ● INDICATES SPACING IMPULSE

Baudot code format

ASCII CODE

The ASCII code, which stands for American Standard Code for Information Interchange, is a 7-level plus parity code that has been established by the computer manufacturing data processing, and terminal manufacturing industries.

Most communication terminals on the market today are designed to conform to ASCII format, however, due to the wide range of applications, some manufacturers make minor changes to the code to make it more applicable to the particular terminal. For example, a card reader line printer terminal that communicates in an asynchronous format might use the SYN code (used for establishing synchronization during synchronous transmission) as a control character to indicate that the printer is out of paper, etc. The parity bit is another change that many manufacturers make. A chart of this code is given in Appendix C.

VARIATIONS OF ASCII CODE

Because ASCII code lends itself to most applications, the majority of manufacturers make every effort to conform to ASCII as much as possible.

One such version of the ASCII code is Data Interchange Code. Primarily, this code differs from ASCII in that some printing characters are replaced by nonprinting control characters, and the parity is specified to be odd. This code, is now readily adaptable to computer-to-computer communication.

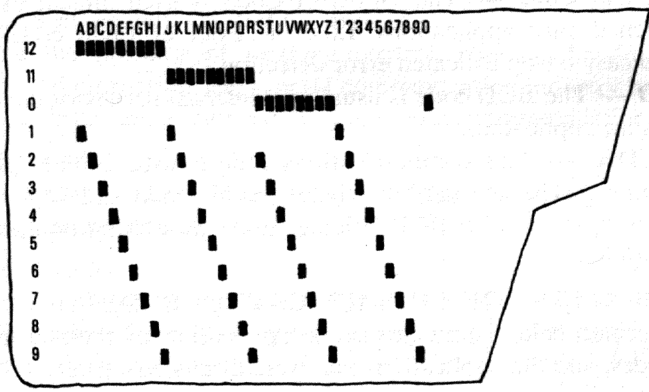
Unlike Baudot Code, the ASCII code has unique code assignments for both alphanumeric and control, and does not require the maintaining of a case status.

FOUR OF EIGHT CODE

The Four of Eight Code is a fixed-ratio code that represents information with a fixed number of one bits and a fixed number of zero bits. In the case of Four-of-Eight, four bits are always one, and four bits are always zeros. Although this code is much less efficient than ASCII, or even Baudot, it is ideal in applications where high accuracy and easy error detection is required. A typical application is credit card verifications.

HOLLERITH CODE

The Hollerith Code is used almost exclusively for punched card applications. It is a 12-level code, designed to represent the alphabet plus digits 1 through 9. This code consists of two parts called zone bits and data bits. There are 3 zone bits and 9 data bits. Only one zone and one data bit are on any one code. Like the Four-of-Eight code, the Hollerith code lends itself to easy detection of errors.



Hollerith code format

BINARY CODED DECIMAL (BCD)

The Binary Coded decimal code is merely a compression of the Hollerith Code. Where the Hollerith code is a fixed ratio code, the BCD code compresses this 12-bit code into a 6-bit code: 2 binary bits replace the 3-zone bits, and 4 binary bits replace the 9-data bits.

EXTENDED BINARY CODED DECIMAL INTERCHANGE CODE (EBCDIC)

The EBCDIC Code is merely the BCD code extended to 8 binary bits. This operation allows the capacity to handle graphic and control character codes. The character capability is very similar to ASCII.

APPLICATIONS OF CODES

The following is an outline of the application of these codes

Baudot — The first serial asynchronous code developed for telegraph keyboard printer and paper tape reader punch communications. It is used exclusively for 5 level telegraph communications.

ASCII and Variations — Used for 8 level telegraph communications, communication control terminals, and computer to computer communication over several communication channels. Most computer and terminal manufacturers conform to the ASCII code, enabling communication to be established between two different manufacturers' pieces of equipment with very little difficulty.

Four of Eight Code — Four of Eight code is used primarily in applications where high reliability is necessary using relatively inexpensive and unsophisticated equipment.

Hollerith Code — The Hollerith Code is used almost exclusively in punched card applications. Like the Four of Eight code, it lends itself to easy unsophisticated error detection.

BCD — The BCD code is usually an internal processor code in data processing application.

EBCDIC — This communications code is used in data processing applications. The compatibility between this code and BCD makes it easily compatible with BCD oriented data processing equipment and applications.

APPLICATIONS OF COMMUNICATION CODES

The chart below summarizes the application of these communication codes, and the applications that were discussed earlier.

APPLICATION	COMMUNICATIONS TECHNIQUE	CODE OR CODES USED
Telegraph and Teletype terminals (Low Speed)	Asynchronous	Baudot 5 level 50 and 75 ASCII 8 level 110 to 15 EBCDIC Used occasionally
Medium speed buffered and unbuffered terminals controllers	Asynchronous Synchronous	ASCII EBCDIC
Intercomputer communications (High Speed)	Synchronous Asynchronous (Used occasionally)	EBCDIC ASCII

Reviewing this chart, we see that with the exception of low speed terminal applications, the codes used are ASCII and EBCDIC. We have also pointed out that although the codes for ASCII characters and EBCDIC are different, the actual code functions are the same.

COMMUNICATIONS USING ASCII AND EBCDIC

At this point we will discuss the application of the ASCII and EBCDIC code as they apply to medium and high-speed bit serial communications between terminal-to-computer and computer-to-computer. It is not the purpose of this discussion to present a detailed description of these formats but to outline how communications are established and maintained as a means of presenting a general understanding of the subject. During the discussion that follows, synchronous techniques will be assumed, for they offer the greatest problems.

Synchronous Communication Control Characters

To overcome the problem of message control, message heading, error detection, and correction, the general communication industry has assigned control characters to be used in synchronous communications.

SYN Synchronous Idle

Used by synchronous transmission systems to provide message and character framing and synchronization.

SOH — Start of Header

Used at the beginning of a sequence of characters to indicate address or routing information. Such a reference is referred to as a Heading. A STX character terminates a heading.

STX — Start of Text

Used at the beginning of a sequence of characters that is to be treated as an entity to reach the ultimate destination. Such a sequence is referred to as a Text. STX may be used to terminate a sequence beginning with SOH.

ETX — End of Text

Used to terminate a sequence of characters started with STX.

ETB — End of Transmission Block

Used to indicate the end of a sequence of characters started with SOH or STX. This ETB character is used when the block structure is not necessarily related to the processing format.

EOT — End of Transmission

Used to indicate a termination of transmission. A transmission may include one or more records and their associated headings.

ACK — Acknowledge

A character sent by the receiving station to the transmitting station to indicate successful reception of a message.

NAK — Negative Acknowledge

A character sent by the receiving station to the transmitting station to indicate unsuccessful reception of a message.

ENQ — Enquiry

A character used to request a response from a remote station. The response that a remote station generates is predefined. Typical response would be, station address, station status content of its buffer, etc.

DLE — Date Link Escape

Used in the text of a message to change the meaning of a limited number of following characters. Typical applications of this character are supplement control information, and transparent text mode of operation. (See Transparent Mode.)

Error — Correction and Detection Characters

In addition to the control characters, which are defined in most formats, one of two error detection characters are specified to be placed at specific points in a transmission. Typically, they immediately follow the ETX or EBT, and their value is derived from characters of text after STX. These two characters are:

LRC — Longitudinal Redundancy Check

The LRC is an accumulated exclusive OR of transmitted characters. The character is generated by the receiving station as data is received and compared with the transmitted LRC. The calculated LRC and the receiving LRC (generated and sent by the transmitting terminal) should be equal.

Cyclic — Redundancy Check

The Cyclic parity check character for longitudinal error control is calculated from a devisor polynomial. A typical one is:

$$X^{16} + X^{15} + X^{2} + 1.$$

TYPICAL MESSAGE FORMATS

The application of synchronous communication techniques throughout the industry are so widespread (and, in many cases, product oriented) that it would be impossible to cover all cases. However, by discussing the two typical facilities of half and full duplex, and the two most common applications terminal-to-computer and computer-to-computer, a basic understanding can be achieved.

Half Duplex Facilities

Quite often there is a tendency by many people to implement full duplex facilities when a half duplex arrangement will do the job. The primary reason for such a decision is because of the great degree of difficulty that half duplex operation causes. However, a half duplex facility is less than the price of a full duplex facility. The main problem that arises when using a half duplex facility is the question of priorities (contention) and the acknowledgement of messages.

The problem is solved in one of two ways, dependent upon application, Master Station and Idle communications mode.

Master Station

This line discipline is used in computer-to-terminal applications where the terminal is slaved to the master, which in this case is the computer. Any transfer that occurs is initiated by the computer. A typical example of such an application would be an alphanumeric CRT display multiplexer communicating with a computer over a synchronous modem.

Most alphanumeric displays buffer the entire text displayed in the CRT: with no computer intervention required, a key on the keyboard (transmit key) provides the means by which the user forwards the text to the computer. However, this key does not usually cause the text to be sent to the computer. It merely gives the text to be transmitted to the computer, when interrogated by the computer.

Typically, when a computer probes the multiplexer it can get one of these responses.

a. Negative Acknowledge
b. No Text Available
c. Text

This mode of operation is ideal for terminal oriented applications where transfers originate at the terminal, and response time is not a significant factor.

Idle Communications Mode

The Idle Communication Mode has two primary advantages over the Master System mode of operation.

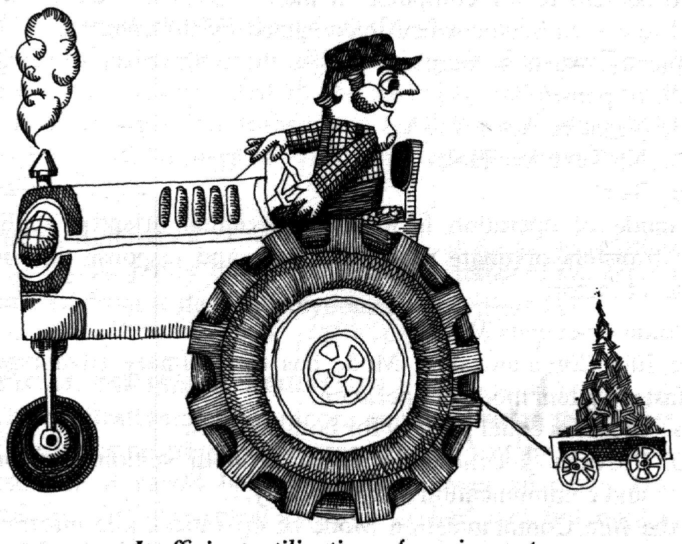
a. Provides equal priority for both stations,
b. Provides a continuous check of both stations' performance and communication link reliability.

In the Idle Communication Mode of operation, idle information is transferred between stations (usually two computers) when text is not being transferred. During this idle mode of operation, data flows from station A to station B for a fixed period of time, and then the flow changes from station B to station A. The only time a station can initiate a transfer is when it is idling the line. The information that is transferred during the idle mode is usually one of three patterns:

a. All ones
b. SYN or some other predefined character
c. Test patterns

Full Duplex Facilities

The main reason for implementing full duplex facilities in a communication system is to increase the data rate. However, in many cases, after the communication (facilities and computer equipment have been installed, and the software running the communication) link turns out to be nothing more than a full-duplex capability being used in the half-duplex mode. Some of the problems that make complete utilization of a full duplex capability difficult are design of available hardware, buffering capability of computer, and message acknowledgement. The hardware and buffering capability problems are a matter of hardware design and core size respectively. The problem of message acknowledgement is a matter of the messages in one direction being of greater length than messages in the other direction. Two typical solutions to the message acknowledgement problem, and a third that is being used in some applications, are outlined below.



Inefficient utilization of equipment

Full Duplex Message Acknowledgement

Insert Acknowledgement — Before forwarding a second message, it is desirable to receive an acknowledgement of the message that the transmitter has just completed transmitting. However, the remote station is transmitting text. To wait for the remote station to complete transmission of text would mean that the transmit facility to the remote station will remain idle until the received message is complete, at which time the message would be acknowledged. This results in inefficient use of a full duplex facility, which is to be avoided when possible.

One approach to the problem is for the remote station to insert an acknowledge or negative acknowledge character in the middle of its transmitted text, and another approach is to interrupt the message with an end of transmission block, insert the acknowledge or negative acknowledge and resume the text with the start of header-start of text sequence. The insertion of acknowledge in the middle of a text is not as reliable as terminating with an end of transmission block because bad character framing can produce a false acknowledgement, and cause an otherwise bad message to be assumed acknowledged and lost. Of course, this is checked with the longitudinal or cyclical redundancy checks, but again this means waiting for the end of the received message and holding up the transmit capability.

The second approach of inserting an acknowledge or negative acknowledge between an end of transmission block, start of header, start of text sequence is more reliable because false characters must not only acknowledge or negative acknowledge to be detected, but it must be preceded by an end of transmission block. The end of transmission block is also followed by a longitudinal or cyclical redundancy check character, but the message does not have to be acknowledged until the end of transmission.

Acknowledgement Numbering — An alternate approach to inserting acknowledgement is to allow a received message to go to completion and continue sending messages with header numbers. When the remote station ends its current transmission, it sends a block of acknowledgement which achieves more efficient use of the facility, but the tradeoff is the loss of buffer area for unacknowledged messages. This is sometimes equally as expensive.

NAK Only — A third approach that is being used is to assume that a message was received properly unless a negative acknowledge character is received within a fixed period of time.

Transparent Mode — The transparent mode of operation is a means by which any bit pattern can be transmitted over the synchronous communications facility. The key to this mode of operation is the DLE character, which is used to indicate this condition.

A good example of this application can be illustrated by idle operation between two computers in which coding or uncoded binary data is being transferred. If in this binary data being transferred, there is a code that is the same as a control character, the message would not be successfully transmitted. However, with the DLE character, uncoded binary data can be transferred. By preceding STX and ETX with a DLE character binary data can be transferred.

The only problem that arises in this mode is transmission of binary codes that are identical to DLE. This is avoided by preceding such a code with a DLE code.

afterword

data communications and the future

Data Communication is a relatively new discipline in the overall field of communication, yet in a brief period of time, as history is reckoned, it has moved into a position of importance. This is due in great part to our technological modern civilization, which is shaping the requirements of data communication systems and is being shaped in turn by the results of data communication. For requirements of virtually instantaneous communications that are imposed by modern technological societies and institutions, the data compression capabilities of digital computer allows the presentation of information in a matter of moments. Thus, digital computer communication holds promise of tremendous potential.

With digital computer equipment and present communication facilities, a complete word-by-word copy of a 27-volume encyclopedia could be transmitted from Maine to California in less than 3 minutes. With this rapid transmission of information in mind, hundreds of American firms are installing digital computer systems to view the results of each day's operations. Stock exchanges, merchandising organizations, moving companies, railroads and airlines are already gaining valuable data. Many of the difficulties that existed a year ago have been surmounted; the difficulties that exist today will be corrected. As an example, in four years, immediate information on firm and individual debits and credits will probably be as close as the digital computer communicator.

glossary of data communications terms

AND Gate	A circuit with multiple inputs that will function only when signals are present at all inputs.
ASCII	American Standard Code for Information Interchange. This is the code established as an American Standard by the American Standards Association.
Alternate Route	A secondary communications path used to reach a destination if the primary path is unavailable.
Audio Frequencies	Frequencies that can be heard by the human ear (usually between 15 and 20,000 Hertz) when transmitted as sound waves.
Automatic Calling Unit (ACU)	A dialing device supplied by the communication common carriers that permits a business machine to dial calls automatically over the communication networks.
Bandwidth	The difference (expressed in Hertz) between the highest and lowest frequencies of a band.
Baseband	In the process of modulation, the baseband is the frequency band occupied by the aggregate of the transmitted signals when first used to modulate a carrier.
Batch Processing	A method of processing in which a number of similar input items are accumulated and processed together.
Baud	A unit of signaling speed. In an equal length code, one baud corresponds to a rate of one signal element per second.

Thus, with a duration of the signal element of 20 ms, the modulation rate is 50 baud

Baudot Code — A code for the transmission of data in which five bits represent one character. It is named for Emile Baudot, a pioneer in printing telegraphy. The name is usually applied to the code used in many teleprinter systems and which was first used by Murray, a contemporary of Baudot.

Bit — A unit of information content. Contraction of "binary digit," a bit is the smallest unit of information in a binary system of notation. It is the choice between two possible states, usually designated one and zero.

Bit Rate — The speed at which bits are transmitted, usually expressed in bits per second.

Block Diagram — A diagram of a system, instrument, computer or program in which selected portions are represented by annotated boxes and interconnecting lines.

Broadcast — The dissemination of information to a number of stations simultaneously.

Buffer — A storage device used to compensate for a difference in rate of flow of data or time of occurrence of events, when transmitting data from one device to another.

Busy Hour — The peak 60-minute period during a business day when the largest volume of communications traffic is handled.

Carriage Return — In a character-by-character printing mechanism, the operation that causes the next character to be printed at the left margin.

Carrier System — A means of obtaining a number of channels over a single path by modulating each channel upon a different "carrier" frequency and demodulating at the receiving point to restore the signals to their original form.

Cathode Ray Tube (CRT) A television-like picture tube used in visual display terminals.

Central Office The place where communications common carriers terminate customer lines and locate the equipment that interconnects those lines.

Channel A path for electrical transmission between two or more points. Also called a circuit, facility, line, link, or path.

Character The actual or coded representation of a digit, letter, or special symbol.

Circuit See Channel.

Code A system of symbols and rules for use in representing information.

Code Conversion The conversion of data from one code to another.

Communications Common Carrier A company that offers its facilities to the public for universal communication services, and which is subject to public utility regulation.

Computer Utility A service that provides computational ability. A "time-shared" computer system. Programs as well as data may be made available to the user. The user also may have his own programs immediately available in the central processor, may have them on call at the computer utility, or he may load them by transmitting them to the computer prior to using them. Certain data and programs are shared by all users of the service; other data and programs because of proprietary nature, have restricted access. Computer utilities are generally accessed by means of data communication subsystems. Also see Service Bureau.

Conversational Mode — A procedure for communication between a terminal and the computer in which each entry from the terminal elicits a response from the computer and vice versa.

Cross Talk — Unwanted insertion of signal from an adjacent communication channel.

Data — Any representations such as characters or analog quantities to which meaning might be assigned.

Data Collection — The act of bringing data from one or more points to a central point.

Data Communications — The movement of encoded information by means of electrical transmission systems.

Data Origination — The earliest stage at which the source material is first put into machine readable form or directly into electrical signals.

DATAPHONE® — A trade mark of the A.T.&.T. Company to identify the data sets manufactured and supplied by the Bell System for use in the transmission of data over the regular telephone network. It is also a service mark of the Bell System that identifies the transmission of data over the regular telephone network (DATAPHONE Service).

Data Processing — Any operation or combination of operations on data.

Data Set — A device that converts the signals of a business machine to signals that are suitable for transmission over communication lines and vice versa. It may also perform other related functions.

Data Transmission — See data communications.

Demodulation — The process of retrieving an original signal from a modulated carrier wave. This technique is used in data sets to make communication signals compatible with business machine signals.

Dial-Up — The use of a dial or push-button telephone to initiate a station-to-station telephone call.

Direct Distance Dialing — A telephone service enabling a user to dial directly to telephones outside the user's local area without aid of an operator.

Display Unit — A device that provides a visual representation of data.

Duplex — In communications, pertaining to a simultaneous two-way and independent transmission in both directions (sometimes referred to as "full duplex"). Contrast with half-duplex.

Error — Any discrepancy between a computed, observed, or measured quantity and the true, specified, or theoretically correct value or condition.

Systematic Error:

A constant error or one that varies in a systematic manner (e.g., equipment misalignment).

Random Error:

An error that varies in a random fashion (e.g., an error resulting from radio static).

Error Control — An arrangement that will detect the presence of errors. In some systems, refinements are added that will correct the detected errors, either by operations on the received data or by transmission from the source.

Exchange — A defined area, served by a communica-

	tions common carrier, within which the carrier furnishes service at the exchange rate and under the regulations applicable in that area as prescribed in the carrier's filed tariffs.
Facility	See Channel.
Facsimile (FAX)	Transmission of pictures, maps, diagrams, etc. The image is scanned at the transmitter, reconstructed at the receiving station, and duplicated on some form of paper.
Flip-Flop	A circuit that is used to store one bit of information.
Foreign Exchange Service	A service that connects a customer's telephone to a remote exchange. This service provides the equivalent of local service from the distant exchange.
Full Duplex	See duplex.
Half-Duplex	Pertaining to an alternate, one-way-at-a-time, independent transmission. Contrast with duplex.
Hard Copy	A printed copy of machine output in readable form, for example, reports, listings, documents, summaries.
Holding Time	The length of time a communication channel is in use for each transmission. Includes both message time and operating time.
Information	The organizational content of a signal.
Information Retrieval	That branch of computer technology concerned with techniques for storing and searching large quantities of data and making selected data available. An information retrieval system may or may not be a real-time system.
In-Plant System	A data handling system confined to one

building or a number of buildings in one locality.

Input
a. The data to be processed.
b. The state or sequence of states occurring on a specified input channel.
c. The device or collective set of devices used for bringing data into another device.
d. A channel for impressing a state on a device or logic element.
e. The process of transferring data from an external storage to an internal storage.

Interface A shared boundary, for example, the boundary between two subsystems or two devices.

Johnson Noise See thermal noise.

Laser A maser that operates at optical frequencies. In communications it acts as a communication device with a coherent optical frequency carrier.

LDX Long Distance Xerography. A name used by the Xerox Corporation to identify its high speed facsimile system. The system uses Xerox terminal equipment and a wideband data communication channel.

Line See Channel.

Line switching The switching technique of temporarily connecting two lines together so that the stations directly exchange information.

Link See Channel.

Local Channel A channel connecting a communication subscriber to a central office.

Maser Acronym for Microwave Amplification by the Stimulated Emission of Radiation.

	A device capable of amplifying or generating radio frequency radiation. Maser amplifiers are used in satellite communication ground stations to amplify the extremely weak signals received from communication satellites.
MICR	Magnetic Ink Character Recognition. Machine recognition of characters printed with magnetic ink. Contrast with OCR.
Message	A communication, prepared for information interchange, in a form suitable for passage through the interchange medium. It includes: a. all portions of the communication such as machine sensible controls, b. an indication of the start of the message and the end of the message, c. a heading containing routing and other information, one or more texts containing the originator-to-addressee communication(s), and the end of text indicator.
Message Format	Rules for the placement of such portions of a message as message heading, address text, and end of message.
Message Numbering	The identifications of each message within a communications system by the assignment of a sequential number.
Message Retrieval	The capability to retrieve a message some time after it has entered an information system.
Message Switching	The switching technique of receiving a message, storing it until the proper outgoing circuit and station are available and then retransmitting it toward its destination.
Microwave	All electromagnetic waves in the radio frequency spectrum above 890 mHertz.

Mnemonic Address — A simple address code that has some easily remembered relationship to the name of the destination; e.g., LA for Los Angeles, ATL for Atlanta.

Modem — Contraction of modulator-demodulator. A device that modulates and demodulates signals transmitted over communication facilities.

Modulation — The process by which some characteristic of a high frequency carrier signal is varied in accordance with another, a lower frequency "information" signal. This technique is used in data sets to make business-machine signals compatible with communication facilities.

Multiple Address Message — A message to be delivered to more than one destination.

Multiplexing — The division of a transmission facility into two or more channels.

Multipoint Circuit — A circuit interconnecting several stations.

Network —

a. A series of points interconnected by communications channels.

b. The switched telephone network is the network of telephone lines normally used for dialed telephone calls.

c. A private network is a network of communications channels confined to the use of one customer.

Noise — In communication theory, an undesired disturbance in a communication system. Noise can generate errors or spurious messages. Contrast with signal.

OCR — Optical Character Recognition. The machine recognition of printed or written characters based on inputs from photoelectric transducers. Contrast with MICR.

Off-Line — Pertaining to equipment or devices not under direct control of the central processing unit. May also be used to describe terminal equipment not connected to a transmission line.

On-Line — Pertaining to equipment or devices in direct communication with the central processing unit. May also be used to describe terminal equipment connected to a transmission line.

One-Way-Channel — A channel that permits transmission in one direction only.

Operating Time — The time required for dialing the call, waiting for the connection to be established, and coordinating the forthcoming transaction with the personnel or equipment at the receiving end.

OR Gate — A circuit with multiple inputs that will function when a signal is present at any input.

Output

a. Data that has been processed.

b. The state or sequence of states occurring on a specified output channel.

c. The device or collective set of devices used for taking data out of a device.

d. A channel for expressing a state of a device or logic element.

e. The process of transferring data from an internal storage to an external storage device.

Parallel Transmission — Method of information transfer in which all bits of a character are sent simultaneously. Contrast with serial transmission.

PICTURE-PHONE ® — A telephone service that permits the user to see as well as talk with the person at the distant end. Trademark of the Bell system.

Page Copy — See hard copy

Path — See channel

Perforator — A keyboard device for punching paper tape.

Print-Out — See hard copy

Polling — A centrally controlled method of calling a number of points to permit them to transmit information.

Priority or Precedence — Controlled transmission of messages in order of their designated importance; e.g., urgent or routine.

Private Line or Private Wire — A channel or circuit furnished a subscriber for his exclusive use.

Punched Paper Tape — A strip of paper on which characters are represented by combinations of punched holes.

Real Time —
a. Pertaining to the actual time during which a physical process takes place.
b. Pertaining to the performance of a computation during a period, short in comparison, with the actual time that the related physical process takes place in order that results of the computations can be used in guiding the physical process.

Redundancy — The portion of the total information contained in a message that can be eliminated without loss of essential information.

Reperforator — A device that automatically punches a paper tape from received signals.

Response Time — The amount of time elapsed between generation of an inquiry at a data communications terminal and receipt of a response at that same terminal. Response time, thus defined, includes: Transmission time to the computer, Processing time at the computer, including access time to obtain any file records needed to answer the inquiry, and transmission time back to the terminal.

Selective Calling The ability of a transmitting station to specify which of several stations on the same line is to receive a message.

Serial Transmission A method of information transfer in which the bits composing a character are sent sequentially. Contrast with parallel transmission.

Service Bureau An installation where the user can lease processing time on a central processor and peripheral equipment. The user supplies the programs and the center will load both program and data to be processed, process the data and deliver the results to the user. The program and data for processing may be delivered or sent between user and center in any of several forms: cards, punched tape, magnetic tape, etc. Data communications may be used between the user and the center to move the information electrically. The service bureau may also provide such services as keypunching the data and preparing it for processing. Also see computer utility.

Signal In communication theory, an intentional disturbance in a communication system. Contrast with noise.

Simplex Channel See one-way channel.

Single-Address Message A message to be delivered to only one destination.

Station One of the input or output points on a communications system.

Status Reports A term used to describe the automatic reports generated by a message-switching system generally covering service condi-

	tions such as circuits and stations out of service and back in service.
Storage	A general term for any device capable of retaining information.
Store-and-Forward	Process of message handling used in a message-switching system.
Stunt-Box	A device to control the nonprinting functions of a teleprinter terminal. Control characters can be sent to it over the communications channel.
Tariff	The published rate for a specific unit of equipment, facility or type of service provided by a communication common carrier.
Telegraphy	A system of communication for the transmission of graphic symbols, usually letters or numerals, by use of a signal code.
Telegraphy, Printing	A method of telegraph operation in which the received signals are automatically recorded in printed characters.
Telemetry	Transmission of data from remote measuring instruments by electrical or (usually) radio means.
Teleprinter	Term used to refer to the equipment used in a printing telegraph system. A teletypewriter.
Teleprinter Exchange Service	A service provided by communication common carriers to interconnect teleprinters. Similar to regular telephone service, customers dial calls from station to station but communicate using teleprinter equipment rather than telephones.
Tele-processing	A form of information handling in which a data processing system utilizes communication facilities.
Teletype	Trademark of the Teletype Corporation.

Usually refers to a series of different types of teleprinter equipment such as transmitters, tape punches, reperforators, and page printers, utilized for communication systems.

Teletypewriter — Term used to refer specifically to teleprint equipment.

Teletypewriter Exchange Service (TWX) — An automatic teleprinter exchange switching service provided by Western Union.

Telex — An automatic teleprinter exchange switching service provided by Western Union.

Telpak — A service offered by communications common carriers for the leasing of wide band channels between two or more points.

Terminal —

a. A point at which information can enter or leave a communication network.

b. An input/output device designed to receive or send source data in an environment associated with the job to be performed and capable of transmitting entries to and obtaining output from the system of which it is a part.

Text — That part of the message which contains the substantive information to be conveyed. Sometimes called "body" of the message.

Thermal Noise — Electromagnetic noise omitted from hot bodies. Sometimes called Johnson Noise.

Tie Line — A private line communication channel of the type provided by communications common carriers for linking two or more points together.

Time-Sharing — A method of operation in which a computer facility is shared by several users for different purposes at (apparently) the same time. Although the computer actually services each user in sequence, the high speed of the computer makes it appear that the users are all handled simultaneously.

Torn-Tape Switching Center — A location where operators tear off incoming printed and punched paper tape and transfer it manually to the proper outgoing circuit.

Touch-Tone — A service mark of the American Telephone and Telegraph Company which identifies its pushbutton dialing service.

TWX — See Teletypewriter Exchange Service

Unattended Operation — The automatic features of a station's operation that permit the transmission and reception of messages on an unattended basis.

Voice Grade Channel — A channel suitable for transmission of speech, digital or analog data, or facsimile, generally with a frequency range of about 300 to 3000 Hertz.

Volatile Display — The nonpermanent image appearing on the screen of a visual display terminal.

WATS — See Wide Area Telephone Service

White Noise — Noise containing all frequencies equally (usually applied to noise containing all frequencies equally in a given bandwidth).

Wide Area Telephone Service — A service provided by telephone companies that permits a customer by use of an access line to make calls to telephones in a specific zone on a dial basis for a flat monthly charge.

Wideband Channel — A channel wider in bandwidth than a voice grade channel.

Word —

a. In telegraphy, six characters (five characters plus one space).

b. In computing, an ordered set of characters that is the normal unit in which information may be stored, transmitted, or operated upon within a computer.

Zipf's Law

An empirical rule that states that the number of times a word appears in a long text is the reciprocal of the order of frequency of occurrence. (For example, the tenth most frequent word will appear 1/10 as many times as the most frequent. Named after George K. Zipf.

appendix a
common carrier facilities

TELEGRAPH FACILITIES AVAILABLE

Common Carrier	Facility Description	Data Rate	Amount of Voice Channel Used	Modem Used
A.T. & T.	Type 1002 Teleprinter Channel	55 Baud	1/12*	None
A.T. & T.	Type 1004 Teleprinter Channel	45 Baud Foreign Exchange	1/12*	None
A.T. & T.	Type 1005 Teleprinter Channel	75 Baud	1/12*	None
A.T. & T.	Type 1006 Teleprinter Channel	150 Baud	1/8*	Bell 100 Series**
Western Union	Class A Teleprinter Channel	50 Baud	1/12	W.U. 1181-A
Western Union	Class B Teleprinter Channel	57 Baud	1/12	W.U. 1181-A
Western Union	Class C Teleprinter Channel	75 Baud	1/12	W.U. 1181-A
Western Union	Class D Teleprinter Channel	180 Baud	1/8	W.U. 1181-A

*NOTE
1/12 and 1/8 may or may not be representative, depending on equipment used.

**Bell System modems are manufactured by Western Electric.

VOICE GRADE FACILITIES AVAILABLE

Common Carrier	Facility Description	Data Rate (Maximum)	Amount of Voice Channel Used	Modem Used	Remarks
A.T. & T.	3002 C1 Voice Channel	2000 Baud	1	Bell 200* Series or Equivalent	Bell Dial Up Network
A.T. & T.	3002 C2 Voice Channel	2400 Baud	1	Bell 200* Series or Equivalent	Private Line
A.T. & T.	3002 C2 Voice Channel	3600 Baud	1	Rixon Sebit-36M or Equivalent	Private Line
A.T. & T.	3002 C4 Voice Channel	4800 Baud	1	Rixon Sebit-48M or Equivalent	Private Line
Western Union	2 KC Broad Band Data Channel	600 Baud	½	W.U. 1601-A or Equivalent	
Western Union	Class E Broad Band Data Channel	1200 Baud	1	W.U. 1601-A or Equivalent	
Western Union	Class F Broad Band Data Channel	2400 Baud	1	W.U. 2121-A or Equivalent	
Western Union	8 KC Broad Band Data Channel	4800 Baud	2	W.U. 2121-A or Equivalent	

*Bell System modems are manufactured by Western Electric.

BROAD BAND FACILITIES AVAILABLE

Common Carrier	Facility Description	Data Rate (Maximum)	Amount of Voice Channel Used	Modem Used
A.T. & T.	Wide Band Channel	19,200 Baud	5	Bell 303B*
A.T. & T.	Wide Band Channel	40,800 Baud	11	Bell 301DS*
A.T. & T.	Wide Band Channel	50,000 Baud	11	Bell 303C*
A.T. & T.	Wide Band Channel	230,400 Baud	58	Bell 303D*
A.T. & T.	Wide Band Channel	200,000 Baud		Bell 303E* Special
Western Union	16 KC Broad Band Channel	9,600	4	W.U. 300 Series
Western Union	48 KC Broad Band Channel	28,000	12	W.U. 300 Series

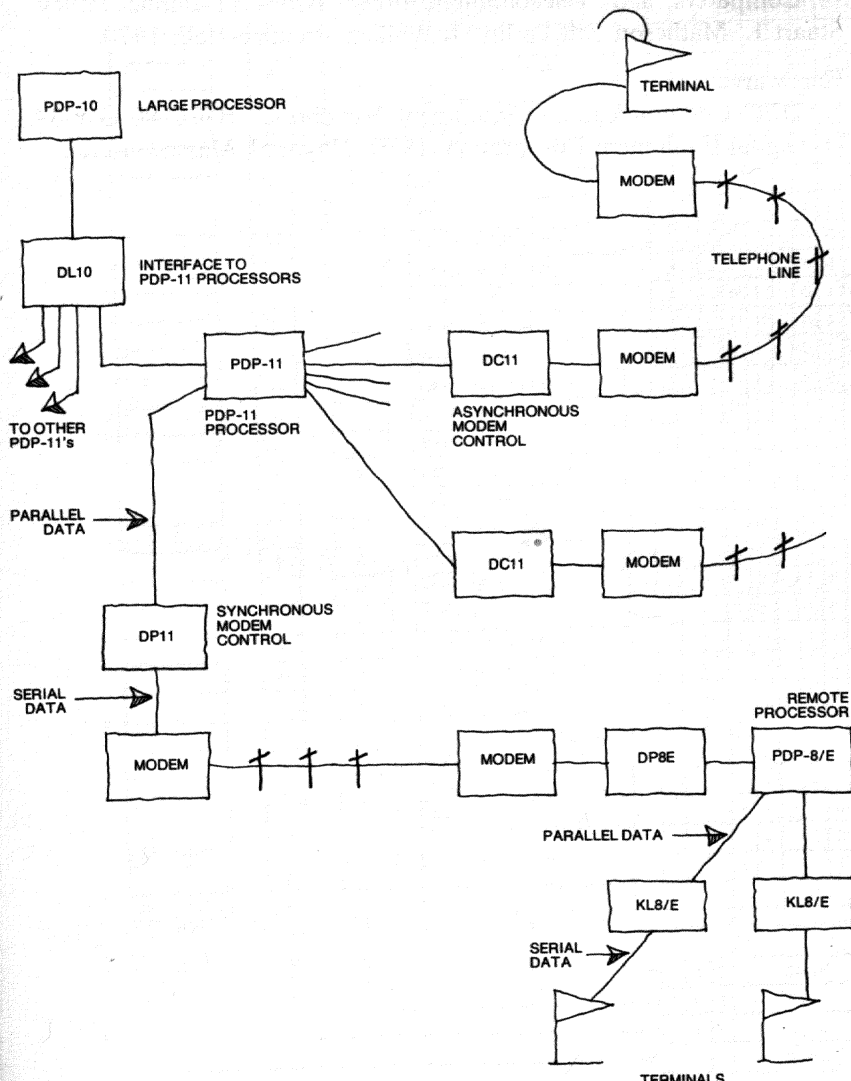
*Bell System modems are manufactured by Western Electric.

REPRESENTATIVE MODEM FACILITIES AVAILABLE

Manufacturer	Model	Speed Maximum	½ or Full Duplex	Sync or Async	Type of Line	Comments
Bell System	103A	300 Baud	Full Duplex	Async	D.D.D.	
Bell System	103E	300 Baud	Full Duplex	Async	D.D.D.	Similar to 103A
Bell System	103F	300 Baud	Full Duplex	Async	Private	
Bell System	113A	300 Baud	Full Duplex	Async	D.D.D.	Originate Only
Bell System	113B	300 Baud	Full Duplex	Async	D.D.D.	Answer Only
Bell System	201A	2000 Baud	Either	Sync	D.D.D.	Full Duplex on 2 calls
Bell System	201B	2400 Baud	Either	Sync	Private	
Bell System	202B	1800 Baud	Either	Async	D.D.D.	
Bell System	202C	1200 Baud	Either	Async	D.D.D.	Full Duplex on 2 calls
Bell System	202D	1800 Baud	Either	Async	Private	
Bell System	205B	600 Baud 1200 Baud 2400 Baud	Full Duplex	Sync	Private	
Bell System	202E Series	1200 Baud	Trans Only	Async	D.D.D. Private	This is a family of modems
Bell System	301B	40,800 Baud	Either	Sync	Private Wide Band	Being Replaced by 303 Series
Bell System	303B, C, D, E	19,000 to 230,400 Baud	Either	Sync	Private Wide Band	
Bell System	811B	110 Baud	Either	Async	TWX Network	
Western Union	118-1A	180			Telegraph	
Western Union	1601-A	600			Voice	
Western Union	2121-A	1200			Voice Broad Band	
Western Union	2241-A	2400	Either	Either	Broad Band	
Western Union	100	200	Either	Async	Voice	
Western Union	200	2400	Either	Async	Voice	
Western Union	300	18,000 40,000	Either	Sync	Broad Band	
Rixon	FM-12	1200	Either	Either	Voice Bell 4A	
Rixon	Sebit 48	4,800	Either	Sync	Voice Bell 4C	
General Electric	TDM 220	2,400	Either	Either	Private Bell 4B	

appendix b

Typical Data Communications System



FURTHER READING

General:

"Telecommunications and the Computer," James Martin, Prentice-Hall, 1969, Englewood Cliffs, New Jersey.

Regulation:

"Computers and Telecommunications: Issues in Public Policy," Stuart L. Matheson and Philip M. Walker, Prentice-Hall, 1970.

Hardware:

"DEC Communications Equipment Handbook" (DEC-00-GW-3A-D) Digital Equipment Corporation, 1970, Maynard, Massachusetts.

appendix C

Model 33 ASR/KSR Teletype Code (ASCII) in Binary Form

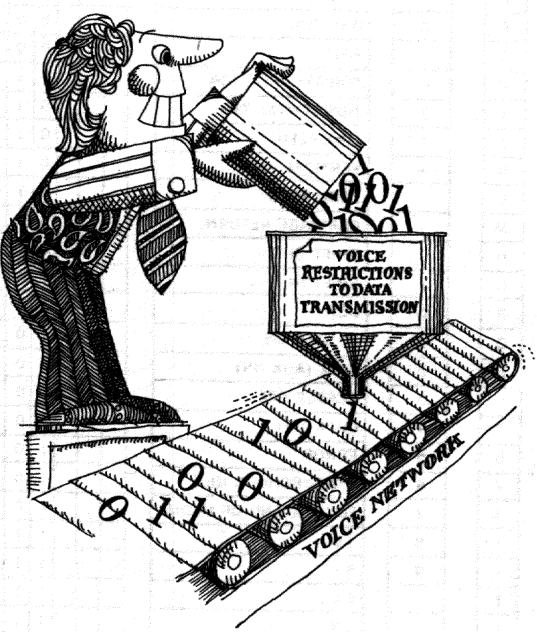
1 = HOLE PUNCHED = MARK
0 = NO HOLE PUNCHED = SPACE

MOST SIGNIFICANT BIT
LEAST SIGNIFICANT BIT

				8	7	6	5	4	S	3	2	1
	@	SPACE	NULL/IDLE				0	0		0	0	0
	A	!	START OF MESSAGE				0	0		0	0	1
	B	"	END OF ADDRESS				0	0		0	1	0
	C	#	END OF MESSAGE				0	0		0	1	1
	D	$	END OF TRANSMISSION				0	0		1	0	0
	E	%	WHO ARE YOU				0	0		1	0	1
	F	&	ARE YOU				0	0		1	1	0
	G	'	BELL				0	0		1	1	1
	H	(	FORMAT EFFECTOR				0	1		0	0	0
	I	)	HORIZONTAL TAB				0	1		0	0	1
	J	*	LINE FEED				0	1		0	1	0
	K	+	VERTICAL TAB				0	1		0	1	1
	L	,	FORM FEED				0	1		1	0	0
	M	—	CARRIAGE RETURN				0	1		1	0	1
	N	.	SHIFT OUT				0	1		1	1	0
	O	/	SHIFT IN				0	1		1	1	1
	P	0	DCO				1	0		0	0	0
	Q	1	READER ON				1	0		0	0	1
	R	2	TAPE (AUX ON)				1	0		0	1	0
	S	3	READER OFF				1	0		0	1	1
	T	4	(AUX OFF)				1	0		1	0	0
	U	5	ERROR				1	0		1	0	1
	V	6	SYNCHRONOUS IDLE				1	0		1	1	0
	W	7	LOGICAL END OF MEDIA				1	0		1	1	1
	X	8	S 0				1	1		0	0	0
	Y	9	S 1				1	1		0	0	1
	Z	:	S 2				1	1		0	1	0
	[	;	S 3				1	1		0	1	1
	\	<	S 4				1	1		1	0	0
Alt-mode	]	=	S 5				1	1		1	0	1
	↑	>	S 6				1	1		1	1	0
RUB OUT	←	?	S 7				1	1		1	1	1

1	0	0	SAME
1	0	1	SAME
1	1	0	SAME
1	1	1	SAME

Note: Bit 8 may be parity or always punched. Always-punched is shown in this table.


VOICE
RESTRICTIONS
TO DATA
TRANSMISSION
VOICE NETWORK
0 1 1
0 0
1 0 1