

MIT OpenCourseWare
<http://ocw.mit.edu>

MAS.632 Speech Interfaces and Mobile Devices
Fall 2008

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.

10

Basics of Telephones

This and the following chapters explain the technology and computer applications of telephones, much as earlier pairs of chapters explored speech coding, synthesis, and recognition. The juxtaposition of telephony with voice processing in a single volume is unusual; what is their relationship? First, the telephone is an ideal means to access interactive voice response services such as those described in Chapter 6. Second, computer-based voice mail will give a strong boost to other uses of stored voice such as those already described in Chapter 4 and to be explored again in Chapter 12. Finally, focusing on *communication*, i.e., the task rather than the technology, leads to a better appreciation of the broad intersection of speech, computers, and our everyday work lives.

This chapter describes the basic telephone operations that transport voice across a network to a remote location. Telephony is changing rapidly in ways that radically modify how we think about and use telephones now and in the future. Conventional telephones are already ubiquitous for the business traveler in industrialized countries, but the rise in personal, portable, wireless telephones is spawning entirely new ways of thinking about universal voice connectivity. The pervasiveness of telephone and computer technologies combined with the critical need to communicate in our professional lives suggest that it would be foolish to ignore the role of the telephone as a speech processing peripheral just like a speaker or microphone.

Although the telephone network was designed for voice communication, it is being increasingly used to carry data either by modem to computer or by facsimile. Technology at either end of the phone call has revealed new roles for voice connections over the telephone network: answering machines and voice mail

allow asynchronous “conversations”; echo cancellation and speakerphones enable multiparty conferences; increasingly flexible call routing services allow more freedom in associating a telephone set with a number; and the telephone accesses a growing number of information retrieval services such as those described in Chapter 6.

It is useful to abstract the major features of a telephone system into the functions of signaling and transport as these concepts transcend the implementation details of any particular network. After many years of relative stability, telephone systems are now changing rapidly with increased use of digital components and radio links to mobile telephones. While network implementation details affect the capabilities of many telephone services provided by computer systems, many of the underlying functional aspects of the network remain unchanged. Analog and digital telephone networks require unique interface hardware and consequently each may present different characteristics of the voice channel. For these reasons, this chapter describes implementations of analog and digital telephone networks in the broader context of the major functional units of any telephone system.

FUNCTIONAL OVERVIEW

This chapter considers telephones from the perspective of their role in providing transport of voice in a system that employs computers for storage, databases, applications, and user interfaces. This view is based on the minimal requirements that make effective use of voice, or any medium as an information type on a computer.

- *Transport:* The ability to move the data from one location to another. Telephones, computer networks, and package delivery services all provide a form of transport.
- *Storage:* The means to save the data for later access. Computer disks hold files, videotapes store images, and answering machines store voice messages.
- *User interface:* A means of allowing a person to retrieve and manipulate the stored information through tools such as editors. The user interface of an answering machine consists of buttons, knobs, and lights.

The telephone network provides the most commonly utilized means of transporting voice, although computer networks have also been used in some experimental systems. Since the telephone network is expansive and eventually will be uniformly digital (i.e., able to transmit speech reliably and with higher quality), researchers choose to use it in many computer scenarios.

We can separate the telephone functions for moving voice between locations into two phases. An **addressing** or **signaling** activity both specifies and establishes a connection to the desired recipient with the help of routing and signaling via the telephone network. Call setup refers to the procedures invoked from the

time the caller begins the call (usually by picking up a handset) until the time the called party answers (again usually by picking up the handset). If a connection is established (i.e., the other party answers), the **transmission** phase begins during which voice is carried over the network. Either party can terminate the call (usually by hanging up), which initiates a final signaling phase of call teardown. Note that if the called party does not answer, a call can proceed directly from the setup to teardown stages. Signaling and transmission phases exist in all telephone systems, but they are implemented differently in analog and digital telephone networks.

ANALOG TELEPHONES

Many of the telephones we use every day including almost all home and public phones are **analog**. As discussed in Chapter 3, analog devices use a continuously varying signal to convey information. Analog phones may be for a single line¹ or for multiple lines with buttons to select between lines and put a call on hold (a **key set**). These phones may have rotary dial or push buttons, which usually generate tones. The analog phone has a ringer (bell) that indicates incoming calls. The telephone is connected to the central office (CO) by wires that run along aboveground poles or in underground cables (see Figure 10.1).

The central office contains the telephone switch, which is connected to the wires from all the telephones in the neighborhood as well as to lines that connect with the remainder of the telephone network. Older telephone switches consist of various types of relays and electromechanical steppers that count the digits as they are dialed, switch lines, and ultimately complete a call. Such switches are rapidly being replaced by modern switches that are digital computers performing essentially the same job as their mechanical predecessors. For each call, the central office connects the telephone to the rest of the network by helping to complete a **circuit** that carries the voice signal between two telephones. A **circuit switch** constructs a path between two points; the circuit is dedicated to this particular path and information on the circuit automatically flows directly to the destination. By contrast, most digital computer networks are **packet switched**, allowing multiple packets of data, each with an embedded address, to share the same physical circuit or wire. The switch characteristics are invisible to the users of the telephone system, however. It is likely that someday all voice telephony will be packet switched although now it is circuit switched.

Analog telephones are connected by a pair of wires to the central office, which is where the first level of telephone switching occurs. Although telephone cable often has four wires, colored red, green, yellow, and black, only the first two are in use in most telephones. The other pair is reserved for special functions such as lighting for illuminated dials, coin collection in a pay phone, or installation of a

¹The classic American telephone from the Western Electric design is called a "2500 set."

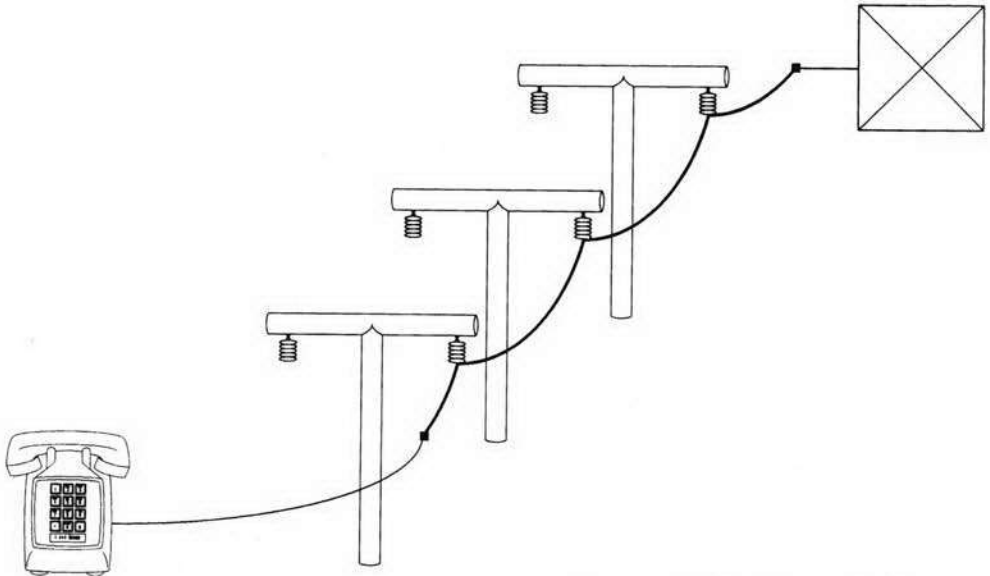


Figure 10.1. A telephone is connected to the central office by a pair of wires.

second line.² The two primary wires, called the **local loop**, are connected to a battery at the CO, which supplies a 48 volt DC voltage. The green wire is connected to the positive side of the battery and is called **tip**, while the red wire, called **ring**, is negative.³ Using a battery to power the loop allows the local telephones to continue operation in the event of a power failure and provides a constant voltage source.

Signaling

Signaling encompasses sending the dialed digits as well as alerting the destination party to an incoming call. Signaling is the transmission of control information party to call management, which includes both **call setup** to establish a voice connection with another party and **call teardown** to free the resources associated with the call at its termination. With analog telephones, signaling is accomplished through a combination of varying the current in the loop and sending audible waveforms over it.

²The specifications discussed in this chapter such as color coding, voltages, frequencies, and durations of audible signals are specific to the United States and may vary in other countries. The discussion of digital telephony is based on international standards.

³These terms date back to the portions of the jacks used to complete connections via patch cords on old manual switchboards.

The wire gauge (thickness) of the wire pair, or loop, at the central office is standardized, and the entire circuit is typically balanced to provide a specified current (20 milliamps) at the telephone. These characteristics provide a uniform operating environment allowing telephones to be attached easily to different circuits. When **on hook**, the telephone presents a high impedance (10 MegaOhms) to the loop and little current flows.

When the handset is picked up, it goes **off hook** and the impedance drops to about 600 Ohms causing a larger amount of current to flow in the loop, which, in turn, signals the switch in the central office that the line seeks attention, and the switch puts a dial tone on the line. The dial tone consists of voltage variations at audible frequencies superimposed on the 48 volt battery; it is transmitted in the same manner as voice. The dial tone notifies the caller that the switch is ready to receive the digits specifying the destination address. The spectral characteristics (i.e., what frequencies are combined in what ratios) of the dial tone allow us to distinguish it from other phone sounds; this is a feedback mechanism designed for human ears.

If the dial tone is heard, the caller may then dial. The verb *dial* refers to the operation of the older style rotary switch telephone (which, incidentally, remains more prevalent in other parts of the world than in the U.S. or Canada). Each digit is entered by rotating the dial a set distance delimited by holes in the dial and a finger stop bar. As the dial is released, it is driven by a spring at a constant angular velocity, and a small switch and cam inside the dial mechanism alternately connect and disconnect the loop wires. The total of such *make/break* cycles specifies a digit. The CO switch detects these pulses as alternate presence and absence of current flowing through the loop. Both the rate of the pulses and the ratio of time during which the loop is open or closed are standardized. The necessity of waiting for the dial to return to its original position before positioning the finger for the next digit enforces a pause between digits.

The switch in the central office counts the digits and makes certain routing decisions locally. In old mechanical switches, the digit pulses actually physically moved relays to close electrical contacts; modern equipment samples the signal on the line and stores the number digitally while deciding how to route it. If the call is local and the first three digits are the same as those of the originating telephone, the call is in the same exchange and the routing of the resulting connection is entirely through the local switch.⁴ If the destination is in the same local area, the local switch may find a free trunk line to the CO switch that serves the dialed exchange. A trunk is a wire pair that may be used to carry calls associated with different numbers for different calls over time as opposed to the wires of a particular telephone, which are dedicated to that telephone exclusively. The local switch will then send the remaining four digits to the remote switch, which is

⁴Again note that this assumes the North American numbering plan, which allows three digits for an area code and seven digits for a number, of which the first three digits specify the exchange. In other parts of the world, different numbers of digits may be allocated to each of these functions.

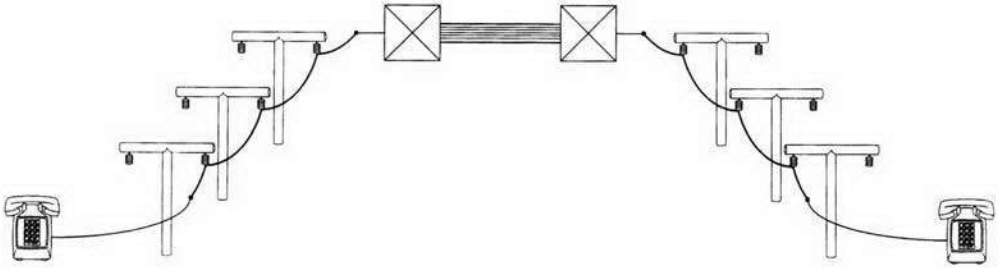


Figure 10.2. Telephones connected through trunk lines between two central offices.

responsible for the next stage of switching, i.e., connecting the trunk to the local loop of the destination telephone, as shown in Figure 10.2.

Switching becomes more complicated in long-distance dialing. In the United States long-distance numbers are usually preceded by a “1” to signal that a 10 digit number follows instead of a 7 digit local number.⁵ *Long-distance* is somewhat of a misnomer these days as the regional operating companies are no longer part of the long-distance company. If two area codes cover adjacent areas, calls may be handled by the same regional company and the service charge may or may not appear as a separate itemized billing item at the end of the month. If the destination is more distant, the call travels through the network of one of the long-distance carriers.

Of course, the call may not be completed. The called party may not answer the phone or the called line may be busy; this information is signaled to the caller by audible tones. These tones, **ringback** and **busy** signals, are distinguishable by the caller because they sound different. Each of the audible signaling tones contains a mix of particular sine waves that compose its spectrum and is presented as an identifiable tone followed by silence. Tones are further differentiated by cadence (the rate at which the tone-silence cycle repeats) and duty cycle (the fraction of each cycle in which sound is present). For example, ringback occurs in six second intervals in the United States with two seconds of sound followed by four seconds of silence.

Other signaling tones convey different information. The **reorder** (or **fast busy**) sound indicates that the network, not the called number, is busy. In the case of a caller error such as dialing an unassigned number or a number that has been changed or not dialing the required “1” for a toll call, a series of three tones in a rise-fall pattern, known as **intercept** (or “special information tones”) is heard, usually followed by a digitized voice explaining the error. All these sounds are indications of **call progress**, i.e., the various steps a call may take towards completion. Call progress tones may be distinguished by differing periods and

⁵ “1” is also the United States’ country code for international dialing so the same number can be used to call across the country or from another country.

duty cycles, and they are designed to be recognized easily by human listeners.⁶ For a computer to recognize the tones it must sample the audio signal from the telephone line to detect cadence, and additionally it may need to analyze the signal's spectrum, which is a more computationally intensive process.

Dialing can also be accomplished by touch tones. When a phone is taken off hook and current flows, depressing a key on the keypad generates a sound, technically called a **DTMF (dual tone multifrequency)** signal. This tone is the sum of two sine waves, one specifying each row and one specifying each column of the keypad. The frequencies were chosen to allow maximal distinction between their harmonics. A decoder in the central office detects the two frequencies and determines which digit was dialed. In addition to the 10 digits, touch tone phones include two extra keys: "*" (asterisk, star) and "#" (sharp, pound sign, number sign, or octothorpe).⁷ Telephone services rely on these keys to differentiate the digits directed to the service from dialed digits specifying a destination number. For example, to cancel variable call forwarding on my telephone I may enter "*73" when I hear the dial tone.⁸

The fact that a telephone has buttons instead of a dial does not necessarily mean it uses DTMF tones to transmit digits. Many less expensive phones have buttons (they are much cheaper to build than dials) but actually use a series of timed make/break pulses like a rotary dial to send the digit. Some of these telephone sets have an option switch between tone and pulse, which is especially useful if a subscriber is not configured for touchtone service at the central office (the local operating company often charges a monthly fee for it) but still wishes to interact with tone-driven interactive telephone services.

The frequencies and amplitude of touch tone signals were chosen to transmit well over the telephone network. When you push a button on your touch tone phone, your phone mutes since the tone generator is quite loud, as someone on another extension will testify. The amplitude makes touch tone detection easy and reliable even over a long-distance connection, which is beneficial for interactive services. It may be difficult to detect DTMF tones while the parties are talking as speech contains energy at a variety of frequencies that interferes with tone detection. Speech may contain frequencies close enough to the touch tone components to trigger false tone detection, especially in less expensive tone detector integrated circuits. This is not a significant issue at the central office, however, as the switch listens only to set up the call, i.e., before any conversation has begun.

⁶There is some variation in the spectrum and cadence of call progress tones. Technical recommendations describe these tones, but network operating companies are not forced to comply. The sounds are consistent enough in the United States that we can usually recognize them but they may be very different abroad.

⁷Actually, the DTMF signaling system supports 16 keys. The extra column, labeled "A" through "D," does not appear on most telephones, however. The extra keys are intended to select a call priority on military telephones.

⁸Of course, I must remember that "73" cancels call forwarding. We will return to the topic of user interfaces to telephone services in the next chapter.

The signaling sounds contain information about the state of the call because they are passed over the same circuit path as the voice signal; they are called **in-band** signaling. Various forms of **out-of-band** signaling may also be provided (e.g., on a key set a light for a particular line will turn on when another extension of that line goes off hook), but out-of-band signaling is mostly limited to digital telephone systems.

When the destination is specified fully by the proper number of dialed digits and the destination line and network are not busy, the destination telephone must audibly alert the called party to the incoming call. This is done by the local switch at the receiving end, which periodically sends a ringing voltage of roughly 100 volts AC down the loop, superimposed on the 48 volts DC. In a traditional phone, this voltage is applied to the windings of an electromagnet; a doorbell-like mechanism then causes a small hammer to vibrate between two bells, producing the classic telephone ringing sound. In more modern phones this signal is detected by a circuit, which makes a sound to imitate the ring using a small transducer. Telephone devices for the deaf detect the ring signal and flash lights for a visible alert. Although this signal as well as the sound of distant ring back that the caller hears are generated in the switch to which the destination phone is connected, the two sounds are often out of phase. This leads to situations where the called party may answer on the first ring, but the calling party has not yet heard any ringback indication.

During the connected portion of the call, the voices of the two parties are transmitted as explained in the next section. Any touch tones generated during this phase of the call are ignored by the switch.⁹ When finished, the parties terminate the call by hanging up their telephones. Going on hook brings the loop back into the high impedance initial state and allows the network to free any resources that may have been associated with the call, i.e., the voice circuit. This termination process is referred to as **tearing down** the call.

Additional in-band signaling may occur during the conversation. For example, a click or special tone indicates an incoming call during a conversation if one party has **call waiting** service. In response, the called party may briefly depress the hook switch (known as a **hook flash**); the switch detects this brief change of line state and puts the first call on hold while completing the second. While on hold, the first call remains connected through the network to the central office associated with the called party, and the circuit resources associated with the call are not freed (nor is billing suspended).

Another new form of signaling information on analog lines is **calling number identification** currently being introduced by regional telephone operating companies. The calling telephone number is transmitted as analog audio in the silent interval after the first ring, using modem-style encoding of digits. Special equipment on the called line can display this number or send it to a local computer.

⁹A few pay phones disable tone generation during the call often to deny callers the option of selecting a long-distance carrier of their choice.

Calling line identification is not yet available on a widespread basis, so the caller's number is generally available only for local calls. Calling number identification is part of a range of new signaling services currently being deployed. These new services include **call blocking** (automatically rejecting calls from a particular number), **distinctive ringing** for calls from a particular number, **call back** (placing a return call to the number that most recently called without the caller knowing what that number was), and **call trace** (which stores the calling number in the switch to be disclosed to law enforcement agencies acting on charges of telephone harassment).

Yet another form of in-band signaling is the **stutter dial tone**, which is associated with voice mail when a message-waiting light is not available. In this situation, the dial tone is not a continuous tone but instead is interrupted, i.e., it stutters, which indicates the presence of new messages in one's voice mailbox. Stutter tone may also be used to indicate an alternate call appearance. For example, an exchange may allow a call to be put on hold by a hook flash. A stutter dial tone then indicates that a new call may be dialed and reminds the caller that there is already a call in progress. After the second call is dialed, the two calls may be conferenced together if desired.

There are many external devices that connect to analog telephone lines, such as answering machines, computer modems, and facsimile machines. These devices employ the same kind of signaling techniques characteristic of analog systems even though some devices transmit data instead of voice. These consumer devices tend to be taken for granted and may be sorely missed in a digital telephone system, which, as described below, uses very different signaling methods.

Transmission

Voice communication can commence once call setup completes. In the simplest case in which both lines are on the same local mechanical switch, the two loops are electrically joined through switching relays and current flows through the completed circuit. Voice is transmitted as current variations in the circuit. The battery from the central office is connected through the local loop to the handset microphone. This microphone typically is full of carbon granules behind a diaphragm located just inside the "mouthpiece." Since sound is composed of variations in air pressure over time, the talker's speech will move the diaphragm, changing the resistance inside the microphone. The loop current changes in proportion to the change in resistance. At the receiving end, the varying loop current flows through a coil around a magnet, which is connected to another diaphragm similar to a conventional loudspeaker. As the diaphragm moves in and out, sound pressure variations are reproduced corresponding to those that impacted the microphone. As a result, we hear the calling party.

The task of the switch and telephone network is to construct a path or circuit for the call. The circuit may actually be a single electrical circuit involving physical connection of the loop coming from the originating phone with the loop terminating in the destination phone. Originally all calls were carried on individual

circuits with patch cords in the central office switchboards completing the circuits. Now it is most likely that once the circuit leaves the local central office, the voice is digitized. The digitized voice is then combined with several other digitized conversations, transmitted some distance by wire, fiber-optic cable, microwaves, or radio transmission to a satellite, eventually demultiplexed from the stream of many conversations, turned back into an analog signal at a distant central office, and sent down the final loop. This entire procedure is virtually transparent to the user except possibly for the delays associated with transmission time on satellite links; we may think of the network as switching virtual analog circuits.

Recall that the loop is a two-wire system. Both parties may talk at the same time and consequently one transmit signal (what one party says) and one receive signal (what the other party says) become mixed on the line. A **hybrid** circuit in the telephone attempts to separate these signals; this is a difficult task in that each connection has a different phase and amplitude response. In other words, the acoustic response of a line varies as a function of frequency, which, in turn, varies from call to call. Some of the signal from the talker's microphone comes back to the talker's earpiece: this is called **sidetone** and provides useful feedback to the caller as a cue to the line quality and a hint as to when to speak up.

But there are several drawbacks to the mix between transmit and receive signals. With telephone-based speech recognition, it is difficult to build an application in which the caller can interrupt ("bargue in") while the computer is talking because the computer's speech will be feeding back into the recognizer. Speaker phones encounter feedback problems when the transmitted signal comes out through the speaker and back into the microphone.¹⁰

Consequently, speaker phones are **half duplex**, meaning that at any moment they are either transmitting or receiving but not both concurrently. Whichever party speaks louder will monopolize the channel ("get the floor"). This makes it much harder to interrupt or offer back channel responses (See chapter 9) and as a result conversation becomes awkward. Several companies (Shure, NEC, VideoTelecom, to name a few) sell full duplex echo-cancellation systems that can be used in a telephone environment. These more sophisticated systems determine the impulse response of the system, i.e., the correspondence between the transmitted and received signal to properly subtract the transmitted element from the received element of the conversation. These devices are still relatively expensive (\$1000 and higher in 1993) pieces of equipment intended for conference rooms, but lower priced consumer models for the desktop are just around the corner.

Telephone audio is band limited to lie between approximately 300 and 3100 Hz. The lower bound is established by blocking capacitors that separate the audio

¹⁰This is not a problem with handsets since the signal coming out of the earpiece is not strong enough to make its way into the mouthpiece. A piece of cotton may often be found inside the handset to prevent acoustic coupling through the hollow body of the handset.

signal from the battery current. The higher frequency is set in part by characteristics of the carbon microphone and in part by the constraints of the digital portions of the network. It is important to limit the upper frequency. The digital segment of the network carries the signal sampled at 8 kHz with 8 bits of μ -law encoded data per sample. As discussed in Chapter 3, μ -law encoding is a logarithmic pulse code modulation scheme, and these 8 bits allow about 12 bits of real dynamic range. The 8 kHz sampling rate cannot represent a signal with frequency components exceeding 4kHz without introducing aliasing distortion; this translates to a little higher than 3kHz for real-world audio filters.

Because of the limited bandwidth some acoustic information is lost during a telephone conversation. A significant amount of low frequency information is lost, generally including the fundamental frequency of voicing, although this does not prevent us from perceiving the pitch of the speech. The higher frequency limit interferes most significantly with the intelligibility of fricatives. These particular frequency limits were chosen with care to allow adequate intelligibility at minimum cost (in terms of bandwidth). Much of the trouble that consumers experience with understanding telephone conversations stems from noisy lines or inadequate amplitude rather than with the bandwidth limitations. Still it is often more difficult to identify a caller over the telephone than in person. On a positive note, users may find synthetic speech less objectionable over the telephone as they are already accustomed to its band-limited characteristics.

Conference calls allow more than two parties to be connected simultaneously. Such a call may be established by a single subscriber with conference service, which allows two calls to be merged by one line; the calls are actually merged in the switch at the request of the common telephone. Local conferencing usually can merge three callers, i.e., two incoming calls; for a larger group a conference service must be involved to establish a **bridge** across the circuits. With a limited number of lines, conferencing can be accomplished by simply adding the signals and perhaps decreasing their amplitude first to prevent clipping speech when multiple parties speak simultaneously. For more than two or three parties, simply adding all the signals is undesirable since the background noise accumulated from each line results in a poor signal-to-noise ratio for the single person speaking. It is better to add together the signals from the one or two lines producing the loudest signals, presumably those conferees who are talking at a given time, and mute the remaining lines to keep their noise out of the conference. This technique interferes with interruption; if many parties speak at once, some will not be heard. Communication is also horribly confounded if inexpensive half-duplex speaker phones are used.

Conference calls are often used as the audio paths for teleconferences in which multiple sites share voice as well as video. Teleconferencing usually requires specially equipped conference rooms. There is growing research interest in less formal conferences in which low cost video equipment is located in ordinary offices, connected together for "video calls," and possibly sharing windows of workstation screens among conferees.

Devices such as modems and facsimile (fax) machines transmit digital data over analog voice circuits. They do so by generating different sounds for "1" bits

and for “0” bits.¹¹ The sounds for various bit patterns are transmitted with differing frequency and/or phase, and the receiver decodes these sounds back into digital data.

DIGITAL TELEPHONES

A digital telephone is one connected to the local switch via a digital connection for both signaling and transmission. Most modern telephone switches are internally digital because digital technologies are often less expensive than analog, digital signaling techniques allow for easier extensions to support new services, and a digital audio signal can be more easily transmitted without degradation.

Unfortunately, all early digital switches used vendor-proprietary signaling techniques, which allowed use of telephones produced by that vendor alone and frustrated attempts to build telephone interfaces for computers. But an international digital standard, ISDN (Integrated Services Digital Network), has been emerging over the past decade, and is the focus of this section. The chief advantage of ISDN over other digital protocols is that it is a standard that ultimately will allow interoperability of equipment across vendors and among countries, making hardware and applications development much more attractive. Unfortunately, true portability across vendors and among countries has not yet been achieved even under ISDN.

With ISDN the connection to each telephone is digital and supports three simultaneous channels called “2B + D.”¹² In such a configuration (Basic Rate ISDN), the 2B, or “bearer,” channels can each carry 64,000 bits of data per second in each direction, which corresponds to the familiar 8 kHz 8 bit μ -law encoding of speech. Speech is usually carried on the B1 channel with the B2 channel reserved for a data call such as the connection of a terminal to a host computer. The 16,000 bits per second D or “data” channel is used for signaling between the phone and the switch. Thus, ISDN signaling is entirely out of band; signaling occurs over a channel entirely different from that which transmits voice.

There are two, three, or four pairs of wires between the switch and the digital telephone depending on how power is supplied to the telephone, but unlike analog phones a pair is not dedicated to each channel. Rather, one pair of wires is for transmit, one is for receive, and the third pair supplies power to the telephone. The bits associated with both B channels and the single D channel are merged together into small data packets (see Figure 10.3), and electrically encoded in a scheme that maintains a balance of “0” and “1” bits on the line.

¹¹Actually most current modem schemes encode pairs of bits, i.e., “00” will be encoded one way and “01” another.

¹²Other configurations are supported. Primary rate ISDN (PRI) supports 23 B channels and a single D channel. Broadband ISDN provides even greater data rates.

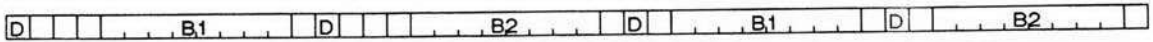


Figure 10.3. The bits of two B channels and one D channel are combined in a single packet.

Signaling

All signaling takes place over the D channel. As with most data communication schemes, ISDN call control may be broken down into a number of layers. The lowest layer, Layer 1, specifies the physical characteristics of the signal including the encoding schemes to balance the bit levels on the wire. Layer 2 defines the addressing scheme, data verification, and retransmission requirements. Layer 3 defines a protocol, Q.931 (ANSI T1.607 in the U.S.), for signaling between the telephone and the switch. Layer 3 is our primary concern for signaling purposes. ISDN currently defines no services above Layer 3. Signaling under ISDN uses data communication schemes similar to those used in computer communication. Although attractive for its power, implementing the ISDN protocol is not trivial, and ISDN telephones (or “terminals” as they are called) or computer interface cards (“network adapters”) are much more expensive than current analog equipment and usually incorporate microprocessors to track the state of each call in progress.

The Layer 2 ISDN D channel protocol is LAP-D (LAP stands for “Link Access Protocol”), or Q.921. LAP-D defines the format of data packets and provides for error checking, retransmission, and flow control. LAP-D also supports a two-tiered addressing scheme to allow multiple data link connections simultaneously over a single D channel. The SAPI (Service Access Point Identifier) portion of the address specifies a service, e.g., a SAPI of 0 identifies the call control service that manages calls on the B channels. Another service, which would use a different SAPI, could be an electronic meter reading service. Digital utility meters could be attached to the ISDN telephone line and accessed from a remote location over the D channel. The second address field is the TEI (Terminal Endpoint Identifier), which corresponds to a particular device being served. Multiple telephone extensions each have a unique TEI so that each can send and receive independent call control messages to the central office switch. For the example meter-reading service, the gas, electric, and water meters might each have its own TEI.

Contained within the LAP-D packets are Layer 3 messages. Layer 3 includes the Q.931 protocol, which implements basic call control. Another portion of Layer 3, Q.932, is designed to implement supplemental services, although the definition of these is still evolving.¹³ Q.931 provides a protocol to set up and tear down voice connections on the B channel as well as data connections on the B or D channels.

¹³Supplementary services include conferencing, call transfer, hold, and drop, as well as managing multiple directory numbers on a single key-set telephone. Most supplementary signaling is at this date switch specific. National ISDN One is an attempt to standardize a core of supplementary services in the U.S.

When a telephone is first plugged in, it brings up its Q.921 layer and establishes a Q.931 link with the switch after negotiating a TEI. When a calling party picks up the handset, a SETUP message is sent to the switch to establish basic parameters such as the type of call; the switch responds with a SETUP ACKnowledgment (see Figure 10.4). Individually dialed digits are then sent as a series of INFO messages to the switch. If all the digits are known in advance, they can be sent within the SETUP message. When enough digits are dialed, the switch returns a CALL PROCEEDING message. Most current ISDN telephones generate audible touch tone frequency sounds during dialing; these tones provide both feedback to the caller and once the call is established in-band signaling support for nonISDN interactive services. During dialing, there may be no B channel established to the switch and the switch does not listen to these tones.

The telephone of the called party then receives its own SETUP message, which may contain INFOrmation fields identifying the number of the calling line as well as other information. The telephone set may use this information to post a message in its display if it has one. The phone itself then generates a ringing sound; this is produced locally not by the central office (although the central office may specify *which* of several different ring patterns to use in support of services requiring distinctive ring sounds). The phone sends back an ALERTing message

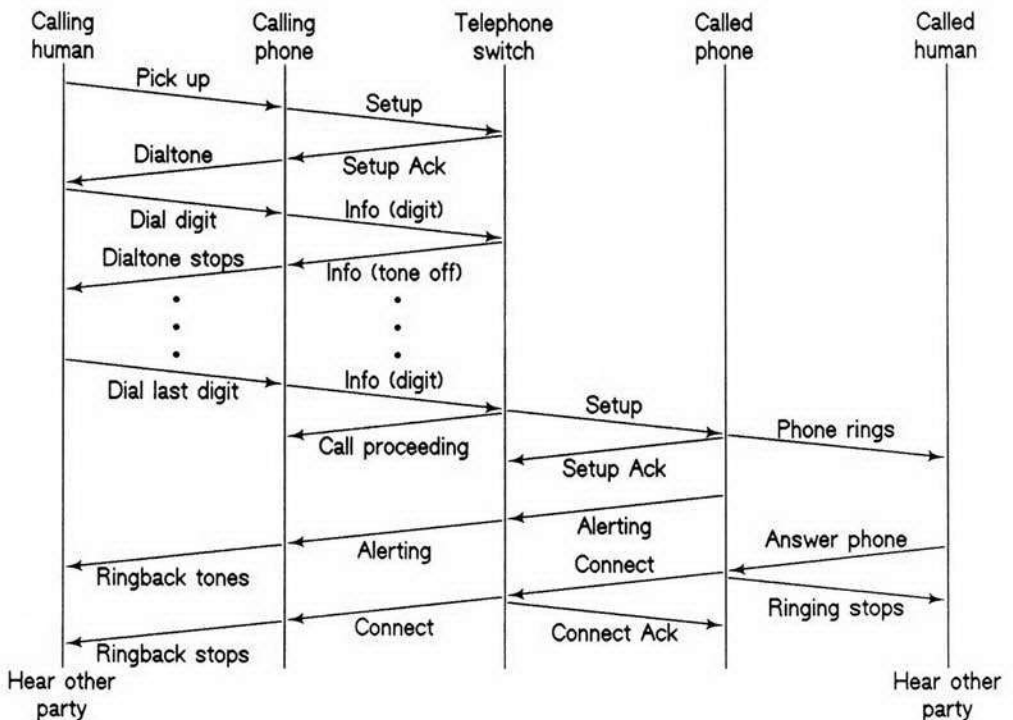


Figure 10.4. The exchange of messages in setting up an ISDN voice call.

to the switch, which passes an ALERTing message back to the calling telephone, which can then generate a ringback indication. If the called party answers the phone, a CONNECT message is sent back to the switch. The switch acknowledges this CONNECT message to the answering party and passes it on to the calling party. At this time, the B channels for each phone are connected and digitized audio is exchanged. At the end of the call, one party hangs up the telephone, which generates a DISCONNECT message, and the switch passes on a RELEASE message to the other phone. Thus, all parties (calling phone, called phone, and both switches) know that the call has been torn down.

Each Layer 3 message is a packet containing a number of fields called information elements. Some information elements are mandatory, some are optional, and, at this point in time, many are specific to a particular switch vendor; these elements are implemented using special "code sets" for each switch. Figure 10.5 shows a sample message and its contents.

Transmission

Digitized voice is carried easily on a B channel. A codec in the telephone digitizes the signal from the microphone as 8 bit μ -law encoded samples at 8 kHz and transmits the samples on the appropriate B channel. At the receiving end, another codec turns the digitally encoded speech into an analog signal, which then drives the earpiece. Applications such as digital answering machines can capture the digital voice data directly from the network to store in a file. An ISDN network is all digital, which allows end-to-end transmission of the digital signal. The main advantage to the consumer is higher quality telephone connections with no difference in audio quality between local and long-distance calls as no additional acoustic information is lost after the initial digitization. Another feature of ISDN voice transmission is that the two sides of a call are carried independently without the mixing that occurs in the analog local loop circuit.

If appropriate equipment is available at each end of the connection, enhanced audio encoding can be carried over the 64,000 bit-per-second B channel. A technique to apply ADPCM coding to several audio frequency sub-bands gives an effective bandwidth of over 7kHz [Mermelstein 1988].¹⁴ The higher bandwidth is especially beneficial for audio conference calls as it facilitates identifying the person talking.

During call setup under ISDN, pressing keypad buttons on the telephone sends digital messages to the switch; any acoustic feedback of standard touch tone sounds is for the convenience of the person dialing. But during the transmission phase of the call, pressing buttons does *not* send messages to the remote switch or telephone. To allow users of interactive voice response services access to those services over ISDN telephones, audible touch tones must be sent in-band over the voice channel. Some digital telephone systems generate the appropriate touch

¹⁴CCITT standard G.722.


```
C SAPI= 0 (CCP) TEI=127
PD=08 UCC Ref=O 10
```

```
M 05 SETUP
```

```
I 04 Bearer Capability          Len=3
    80 Coding standard          CCITT
        Transfer capability      Speech
    90 Transfer mode            Circuit
        Transfer rate            64 kbits/s
A2 Layer 1 protocol            G.711 u-law
I 18 Channel Identification     Len=1
    88 Interface identifier explicit No
        Interface type          Basic
        Indicated channel exclusive Yes
        Channel is D-channel     No
        Channel selection        None
I 1E Progress indicator        Len=2
    82 Location                  Local public
    83 Progress description     Orig non-ISDN
I 34 Signal                     Len=1
    4F Alerting off
```

```
C SAPI= 0 (CCP) TEI= 64
```

```
PD=08 UCC Ref=D 10
```

```
M 01 ALERTing
```

```
I 18 Channel Identification     Len=1
    81 Interface identifier explicit No
        Interface type          Basic
        Indicated channel exclusive No
        Channel is D-channel     No
        Channel selection        B1
```

Figure 10.5. The information elements contained in a SETUP message and the ALERTing message sent in response.

tone sounds in the local switch instead of the telephone when the caller uses the keypad, but this provides less effective auditory feedback to the user.

PBXs

Not all telephones are connected directly to the central office. In business environments with a large number of telephone lines, it is more common that the phone in an office is connected to a switch owned by the business called a PBX (Private Branch eXchange). The PBX provides an initial level of switching internal to the organization. As illustrated in Figure 10.6, a PBX can either connect two internal phones to each other (without incurring a charge from the telephone operating company as none of its resources are used) or connect an internal telephone to an external number through trunks to the central office. In an alternative arrangement sold under the trademark Centrex, each telephone is connected to the central office, but software in the central office switch emulates the internal number dialing and other features of an independent PBX.

PBXs provided many calling features, such as conference, transfer, and hold long before these were available to single line customers.¹⁵ Because PBXs usually serve many fewer lines than a central office, they represent a much smaller capital investment and are marketed by many vendors. Because their interconnection with the public switched telephone network is limited and well specified, the PBX is free to implement its own internal call management protocol. Presently PBX vendors offer a variety of incompatible and proprietary digital protocols to support services such as display of calling party and call routing information such as "forwarded from" that can be used by a receptionist answering many lines.

PBXs also offer the flexibility of providing their own numbering plan. For example, in the sample PBX in Figure 10.6 all the internal extensions consist of two digits beginning with "2." This numbering plan interferes with dialing the outside number 232-1234 as the internal call to extension "23" would be initiated when the second digit of that number was dialed. Although the PBX could delay after the second digit waiting to see if the caller dials any additional digits, this would slow down all internal calls. Instead, most PBXs employ a numbering plan specifying that the prefix "9" precede all outside calls.

PBXs are often used to form private telephone networks within large companies that are geographically dispersed. For example, the PBX we have been using as an example might provide calling to a remote office when the first dialed digit is "8"; to reach telephone "25" at the remote site, the caller would dial "825." The PBX would route this call to a trunk (see Figure 10.7) leased from one of the long-distance operating companies tied to the distant PBX. The distant PBX would then complete the circuit to the local telephone line numbered "25." This archi-

¹⁵A set of central office services called CLASS (Custom Local Area Signaling Services) provide these as well as many of the other features discussed in this chapter, but CLASS is only now becoming widely available in the U.S.

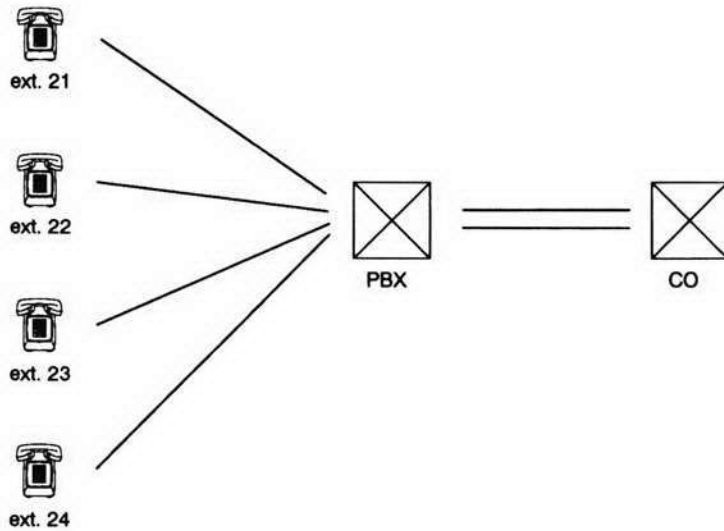


Figure 10.6 A PBX provides an internal telephone network. Telephones served directly by the PBX can be connected without using the public telephone network. The PBX can connect an internal telephone to the central office switch through trunk lines.

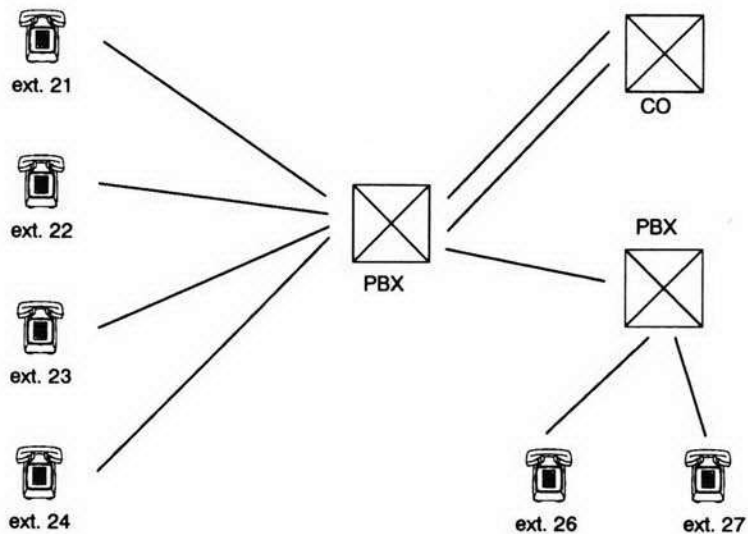


Figure 10.7 A PBX may have trunks to other PBXs as well as the central office, providing private telephone networks.

ture mirrors the arrangement of central offices and connections to remote central offices through the public telephone network; the only difference is who owns the equipment.

In practice it is more common that each telephone number is unique. The PBX would know from its routing tables that telephone numbers in the range of "25" to "30" are at the remote site and route these calls to the distant switch. In any case, the digits "25" must be transmitted to the remote switch so that it can route the call to the appropriate remote telephone.

SUMMARY

The telephone plays an important role in many aspects of using computers for communication. Telephone-based interactive voice services provide remote access to a variety of databases. The next chapter discusses computers as mediators of telephone conversations between people. Chapter 12 describes stored voice as a computer data type; at the moment, most such data arrives over the telephone, e.g., voice mail messages. This chapter has discussed operational characteristics of telephone networks so that the reader appreciates both the underlying issues in using computers to control calls and how properties of the telephone voice channel require particular interaction techniques for voice services.

Telephone management can be broken down into a signaling phase that sets up and tears down calls and a transmission phase during which two or more parties converse. Signaling and transmission are handled very differently in analog and digital networks, although most *functionality* can be had using either. Analog signaling is in-band, i.e., carried as audible sounds. A computer interfaced to the analog network must listen in and analyze the acoustic properties of sounds on the local loop to track the progress of a call or use hardware that does essentially the same thing. Digital signaling uses protocols similar to those used in existing computer networks and hence can be more easily integrated with other workstation operating system activity than analog signaling. Digital signaling is also faster and allows more readily for management of multiple calls simultaneously.

Analog and basic digital voice transmission support similar audio coding quality and bandwidth; the data rate of the ISDN B channel was chosen so as to model the characteristics of analog telephone circuits. But the digital network offers several additional advantages. Foremost in an all-digital network there are no further reductions in audio quality after digitization no matter how many miles the circuit spans; analog circuits exhibit signal degradation at each stage of amplification, which can be numerous over a very long distance. An end-to-end ISDN circuit keeps both parties' voices separate, making it easier to use speech recognition to interrupt an interactive service while it is speaking. The digital network can be used to carry speech based on improved coders offering higher audio bandwidth over existing channel capacity. As ISDN offers multiple B channels on a single telephone line, voice and data call or multiple voice calls may be in progress simultaneously over a single connection to the network.

Although ISDN is attractive for many reasons, it will be slow to penetrate the consumer market because digital telephone equipment is more expensive than analog, and few residential customers require telecommunication services that can benefit from ISDN's capabilities. At least in North America, the first wave of ISDN deployment has been limited generally to select business environments. Although the economics of wide-area networking or the need to merge voice and data services dominate deployment decisions, business environments also have the most to gain from integration of computers and telephones for call management—which brings us to the topic of the next chapter.

FURTHER READING

There are a limited number of references that explain telecommunications technologies to a broad audience. Chorafas presents a survey of network technologies and weaves them into their roles in an evolving global network. Stallings presents ISDN in a very readable manner. Briley focuses on details of switching in the telephone network, while Talley concentrates on data communication protocols; both books are readable by a general audience but more focused than was this chapter.