

# INTERACTIVE VOICE TECHNOLOGY APPLICATIONS

David R. Fischell, Sarbmeet S. Kanwal, and Daniel Furman

**David R. Fischell** is supervisor of the Enhanced Services Complex Advanced Development Group in West Long Branch, New Jersey. **Daniel Furman** is head of the Consumer Services and Technologies Department in Murray Hill, New Jersey. **Sarbmeet S. Kanwal** is a supervisor in the Voice Transaction Systems Department in Columbus, Ohio. They are with AT&T Bell Laboratories. Mr. Fischell is responsible for taking feasible concepts for interactive-voice services from prototype development, systems requirements, and technical trials to service availability. He joined the company in 1979 and has a B.S. in engineering physics, an M.S. in applied physics, and a Ph.D. in applied physics and electrical engineering, all from Cornell University (Ithaca, New York). Mr. Furman is responsible for applied research on marketing (continued on page 76)

Voice, data, and video information can be delivered to our homes or businesses using today's technology. But the U.S. market for multimedia information services is in its infancy, and interactive-voice services are merely the first step on the evolutionary path. Applications that use interactive voice fall into four categories: information delivery, entertainment, voice messaging, and transactions. All share a common trait—a person interacts with a machine over telephone lines by sending Touch-Tone signals in response to stored voice messages. AT&T's interactive-voice service offerings are based on a common network platform—the *Enhanced Services Complex* (ESC). An ESC is a collection of switching, speech processing, and computer equipment networked together to provide enhanced voice and data services over the public switched telecommunications network. Currently, these services enable callers to interact with and manipulate stored audio information. Soon, the ESC will include capabilities for speaker verification, speech recognition, and facsimile transmission.

## Information via Telecommunication

Today's technology can deliver information services to our homes using voice, data, or video as the medium. However, the market for such services in the United States has been slow to develop. Why has it been so slow? And, what will it take to make this market develop? These important questions confront all providers of information and telecommunications services.

To address these questions, we must begin with the three components required for a new information service to be successful:

- Consumers must be willing to pay for the service, including the costs of leasing or purchasing the terminal equipment required and

**Panel 1. Acronyms in this Paper**

ADPCM	adaptive-differential pulse-code modulation
ASN	AT&T switched network
ATM	automated-teller machine
AUDIX	audio-information exchange
CEI	comparably efficient interconnect
DCIU	data communications interface unit
DCP	digital communication protocol
DDS	digital data service
ESC	Enhanced Services Complex
ISDN	Integrated Services Digital Network
LEC	local-exchange carrier
OA&M	operations administration and maintenance
PBX	private branch exchange
PC	personal computer
PRI	primary-rate interface
RBOC	Regional Bell Operating Company
SDN	software-defined network
T1	1.544 Mb/s digital trunks with 24 voice channels
VRU	voice-response unit
WGS	WorkGroup System

of leasing telecommunications facilities.

- The information provider must be willing to pay the production costs of assembling the information for a new medium.
- The telecommunications company (including terminal equipment manufacturers) must be willing to invest in the facilities required to deliver these services.

This set of circumstances leads to a classic *chicken and egg* situation: any provider who moves ahead faces significant risk if the other two components do not follow suit. Hence, we have the *information-service triangle* (Figure 1) where, to move forward, all parties must be willing to invest.

What if all parties are not willing? *Government intervention* is one solution, as was the case for Minitel in France. The French Government, which owns the

telephone network, gave a Minitel videotex terminal, free of charge, to each household. *Cross ownership* is another possibility. RCA, which built the first television sets, owned NBC and could, therefore, go from national radio to national television with reduced risk. But in the absence of government intervention or cross ownership, an *evolutionary approach* is the one most likely to succeed in the U.S. marketplace.

The evolutionary path for any new service is a series of incremental steps that develop the extensive capital infrastructure required to build the expensive production and transmission facilities, and educate consumers in the use of unfamiliar technology. One example of this is the evolution from broadcast TV to cable TV, to premium-pay TV, and now to *pay-per-view* TV.<sup>1</sup> (Panel 1 defines acronyms used in this paper.)

Obviously, interactive-voice services are the logical first step to begin the evolution to multimedia information services in the U.S. market. The terminal equipment required for interactive-voice services (i.e., the telephone) is universally available in American homes, and the national public telephone network makes the provisioning of special transmission facilities unnecessary. Thus, two of the three components required for information services have a greatly reduced risk factor for voice-information services, which increases the chances for success.

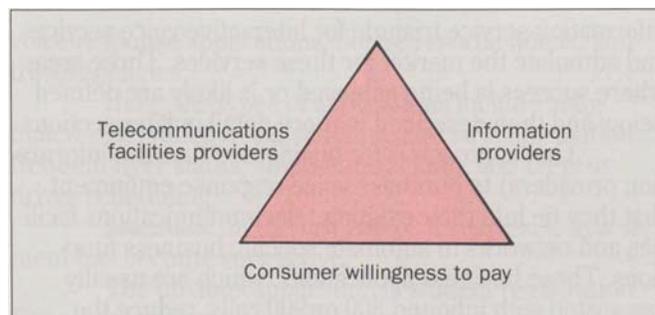
### **Interactive-Voice Services**

Passive-voice-information services, i.e., 976-type services, have been offered successfully for more than 20 years now. However, these services are limited to the playback of a single, recorded, voice message. Because of the wide penetration of Touch-Tone service and availability of affordable speech-processing technology based on mass-produced PC technologies, passive-voice services are rapidly being replaced by a much richer array of interactive-voice services. The promise of robust and inexpensive speech-recognition technology in the near future will help to extend the reach of these services to

rotary-dial users and will make these services more appealing to the general public. (By *robust*, we mean the technology is reliable and extensive.)

**What Are These Services?** Interactive-voice services are generically those that involve a person interacting with a machine over a telephone line by sending Touch-Tone signals and receiving recorded voice responses. Applications of these services fall into one of four categories:

- **Information distribution.** The retrieval of stored-voice programming. This includes:
  - Traditional *audiotex* applications such as news, weather, and sports that play back the latest recorded message.
  - Information services that involve simple databases. An example is a train schedule and fare application that allows callers to select the originating station, termination station, departure date, and time. The service then plays back the schedule and fares for trains that meet the input criteria.
- **Entertainment.** This is the fastest growing area of the interactive-voice services due to the availability of 900-number services that pay the information provider for each call generated. These applications include such services as horoscopes and soap-opera updates, as well as games and contests with prizes. Through advertising, information providers can create huge call volumes.
- **Voice messaging.** The delivery and retrieval of stored-voice messages. Examples include:
  - Network voice-mail services. An individual is assigned a voice mailbox where others can leave messages that the mailbox owner can retrieve.
  - Call-delivery services. Anyone can leave a message to be delivered automatically to a second party.
- **Transactions.** This type of service involves both the retrieval and the updating or modification of stored information. The category includes financial services and fundamental business functions such as order entry, order tracking, dealer locator, and customer



**Figure 1. Information-service triangle. Consumers must be willing to pay for the service and the equipment needed. Information providers must invest in assembling the content for the medium. Telecommunications companies must provide the facilities needed to deliver the services.**

service. For example, the Schwab TeleBroker Service<sup>SM</sup> discount stock-brokerage service allows callers to obtain current stock prices and buy and sell securities using only a Touch-Tone telephone. (TeleBroker Service is a service mark of Charles Schwab & Co., Inc.) The service facilities interact both with the caller and, over a data link, with the brokerage firm's host database to complete a transaction.

The common element among these applications is the technology: a terminal (the Touch-Tone telephone), a telecommunications network, a voice-response unit (VRU), and an information source. The consumer dials into the voice-response unit over the telecommunications network. Then, he or she interacts with the voice-response unit—using Touch-Tone input and voice response—to obtain the desired information, leave or retrieve a voice message, or perform a transaction.

The information may reside in the voice-response unit itself or may be retrieved from an external database. An option to transfer the call to a human attendant or agent may also be present.

**Ways to Stimulate the Market.** Currently, industry leaders are trying several approaches to break the

---

information-service triangle for interactive-voice services and stimulate the market for these services. Three areas where success is being achieved or is likely are defined below and then described in more detail in later sections.

One approach is for businesses (i.e., the information providers) to purchase voice-response equipment that they tie into their existing telecommunications facilities and networks to automate specific business functions. These business applications, which are usually associated with inbound 800 or 900 calls, reduce the number of telephone agents needed to screen incoming calls. Hence, a company can save money by reassigning its agents to more important tasks. Also, these applications improve customer service (i.e., 24-hour availability), which can give the company a competitive edge.

Other companies use interactive-voice services as an innovative advertising medium. To attract new customers, they offer giveaways and contest prizes through telephone calls. (For example, anyone who calls the 800 or 900 number in the advertisement must listen to product information to receive a prize.) Although this approach is growing, its major drawback for the businesses involved is the high cost of capital associated with significant use of this technology.

Another approach that will be growing in the future is cross ownership, which is being tried mainly by the local and interexchange telephone companies. It is their response to some recent relief in the regulation of enhanced services that do not require an information provider, such as voice-messaging services. These companies are also trying electronic publishing, where the company owns or creates the information. In both cases, the information-services triangle is broken because the terminal equipment is already in place.

The third approach is the one that is having the best success. Here, a company—usually a long-distance carrier or a large service bureau—sets up large enhanced-services facilities to try to create a capital infrastructure for interactive-voice services. The long-distance carrier

has the network know-how, and the large service bureaus have the agents and telemarketing experience.

This is an example of the evolutionary approach, where the provider with the enhanced-services facility sells its capacity to the information provider or another business on a usage-fee basis. Thus, the originator of the service greatly reduces the up-front risk and can concentrate on information creation and service promotion. This is attractive for both the caller-paid (i.e., 900-number) services offered by information providers for profit, and the free-to-callers (i.e., 800-number) services offered by businesses interested in cost savings (i.e., the business needs fewer telephone agents to handle the calls).

#### **AT&T's Voice-Response Equipment**

AT&T is one of the leading vendors of voice-response equipment in the marketplace. AT&T manufactures and sells the Conversant® voice-information system and AUDIX Voice Mail system. (AUDIX stands for *audio information exchange*.)

The Conversant voice-information system is an AT&T 6386 WorkGroup System (WGS) computer that runs the UNIX® operating system and has special voice hardware and software to allow up to 48 channels of interactive-voice capability. (UNIX is a registered trademark of UNIX System Laboratories, Inc.) AUDIX is AT&T's premier voice-mail system designed to work with AT&T's Definity® telecommunications systems and other PBX central-office switches.

The Conversant and AUDIX systems are fully described in another paper in this issue.<sup>2</sup>

#### **Business-Related Applications**

These interactive-voice service applications are usually aimed at enhancing some aspect of the operation of a business or an institution. As pointed out earlier, these services do not need special terminal equipment and usually do not require the caller to pay a fee. Hence, the investment decision is all within one company. Most

---

medium to large companies can justify the expenditure of capital for VRU equipment that has a payback period of less than one year.

In the rest of this section, we identify some of the major industry segments that have been active in offering interactive-voice services to their customers, as well as their employees. For these applications, some businesses also purchase the equipment; others just purchase the service from telecommunications providers.

**Retail and Wholesale Sales.** Automated order entry is a common application in the sales industry for catalog orders on the retail side and for restocking inventory on the wholesale side. A caller uses his or her Touch-Tone dial to interact with a voice-response unit, which is connected to the host computer. The computer contains product-inventory information and can generate paper orders without involving a human agent.

Other applications in the industry include customer-account inquiry, order-status inquiry, and delivery scheduling. These are transaction-type services.

**Banking and Finance.** This industry segment was a leader in the use of voice-response technology to improve customer service and reduce costs (i.e., fewer human agents are needed). The transaction services that are popular with banks are: account-balance inquiry, mortgage-interest-rate inquiry, money transactions between accounts, credit-card authorization, locator service for automated teller machines (ATMs), and check clearance.

In the finance or investment industry, services such as investor-account information, mutual-fund and stock quotes, and financial-trends inquiry are popular.

**Insurance.** The insurance industry is another large user of interactive-voice services. Claim reporting and tracking is one of the main applications here, but others—such as agent locator, claim-office locator, and application prequalification (for coverage or claims)—are also offered.

**Transportation and Travel.** Airlines, railroads, and trucking companies are potential heavy users of

voice-response applications. So are resorts, hotels, and travel agencies.

Here, the types of applications include automated reservation desk, schedule and fare information, frequent-flyer status, shipment tracking, and crew or driver scheduling.

**Education.** In the last two or three years, this segment has become very active.

The obvious application is student registration. But other applications such as fee payment, information services (e.g., course content, campus activities, housing availability, financial aid), and ticket purchasing (e.g., tickets for sports events, concerts, plays) are also in use.

**Government Agencies.** Government agencies use many agents just to disseminate information to the public. Their agents may answer questions, take orders for documents, or direct the call to the right office of the agency. Interactive-voice services can help these agencies save money and provide better service. The Internal Revenue Service, U.S. Postal Service, Social Security agencies, and state motor-vehicle and lottery agencies are all good candidates for offering information and transaction services.

#### **Telecommunications-Company-Owned Services**

The telecommunications companies [i.e., local-exchange carriers (LECs) and long-distance or interexchange carriers] are in a good position to offer new information services that use the existing public network and are integrated with it to varying degrees. The consumer franchise that these companies enjoy—with a reach into nearly every household—and each company's ability to raise the capital needed for large-scale transmission and VRU facilities are difficult to match.

AT&T and all seven Regional Bell Operating Companies (RBOCs) have begun to offer these services where regulation permits. The services include *contentless* applications, such as voice messaging, and *content* applications with information providers.

---

**Messaging Services.** These were the first type of enhanced services to be allowed through the regulatory relief. Since 1987, the RBOCs have been permitted to offer voice-messaging services, which temporarily store a customer's message for retrieval by or delivery to the person for whom the message was recorded. All seven RBOCs have already offered voice-messaging services or have them under trial. These services represent an assortment from *call answer* (which functions as an answering machine) to *message delivery*.

AT&T has recently announced two voice-messaging services:

- AT&T Voice Mail to provide network voice mailboxes with store-and-forward capabilities. (AT&T Voice Mail is described in the applications section of this paper.)
- A message-delivery service called AT&T VoiceMark® messaging service.<sup>3</sup>

AT&T VoiceMark messaging service is a new service that permits a caller to record messages in his or her own voice to be delivered at times and to locations (i.e., telephone numbers) that he or she chooses. VoiceMark messaging service offers many benefits. The service does not require special equipment (other than a Touch-Tone telephone), and neither senders nor recipients need to presubscribe to the service. Thus, any sender can communicate with any recipient, and the recipient does not need to take any special action (such as message retrieval). Senders can also communicate with recipients even when they may not be available at the same time. Hence, this can be an effective time-management tool, especially when senders and recipients are in different time zones.

Among the major features of VoiceMark messaging service is ubiquitous availability. To send a message, the caller may access the service from any Touch-Tone telephone in the United States by dialing a single toll-free 800 number. From foreign countries, he or she only needs to use AT&T's USA Direct® telecommunications service. In addition, VoiceMark messaging service may

be accessed via AT&T High Seas Radio Service.

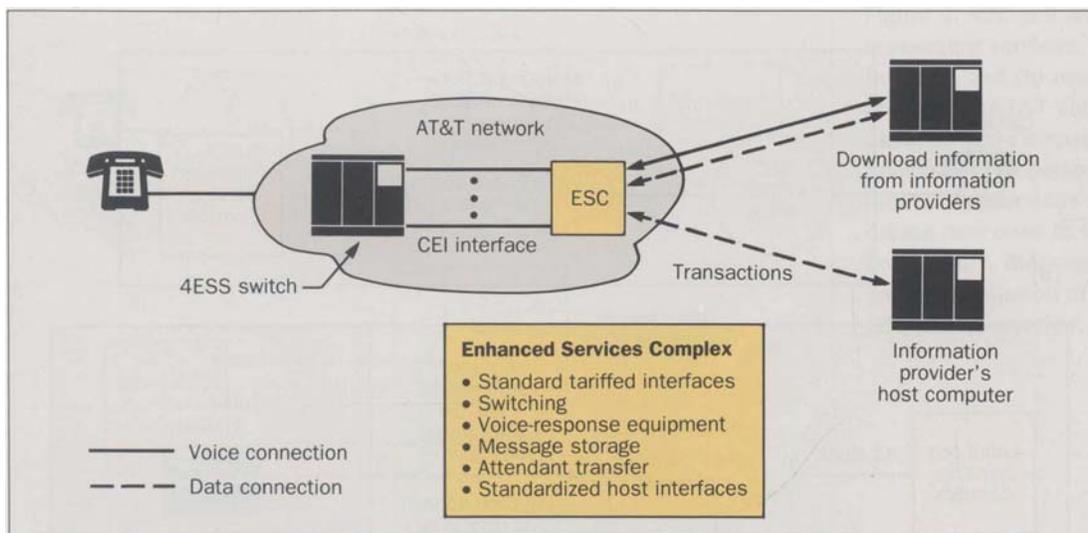
Messages are recorded and delivered in the sender's voice, and the sender specifies the delivery location (international, as well as national) and the date and time of delivery.

**Information Services.** Because current regulations prohibit the RBOCs from offering electronic publishing (i.e., giving customers access to RBOC-owned information databases), few services in this information area are offered in the local arena. One service allowed is the Talking Yellow Pages® directory service, where a caller can listen to recorded information about a business after dialing a telephone number and entering an identification number provided in the business' Yellow Pages advertisement. (Talking Yellow Pages is a registered trademark of Reuben H. Donnelly.)

Another service under consideration is *reverse directory*, where a caller enters a telephone number and hears the name and address of the associated subscriber. With this service, a person who subscribes to an RBOC's caller identification service could find out whose number is displayed on the call identification unit. Also, law enforcement agencies and fire departments who, during an emergency call, were given only a telephone number could determine the location of the emergency.

**Transaction Services.** Touch-Tone signals and voice response have traditionally been used for transaction services, such as call routing and calling-card validation. As more network features are provided, it becomes advantageous to the service providers to use interactive voice to enhance these features by providing security, i.e., by allowing entry and validation of authorization codes.

The network remote-access feature of AT&T's software-defined network (SDN)<sup>4</sup> uses interactive-voice capabilities to permit business callers to access their private networks and make sequential calls. (SDN is based on the public switched network and offers customer-defined capabilities that reside in databases at network control points.) Other transaction services that are not so closely



**Figure 2. AT&T's Enhanced Services Complex (ESC) for interactive-voice services is a common platform for provisioning enhanced network services, i.e., services that involve the storage and retrieval of information or the conversion of information protocols. AT&T is permitted to**

related to the network function are also being considered by AT&T and other telecommunications companies. For example, AT&T currently has in trial a home-banking service that will allow subscribers to pay their monthly bills by executing transactions on their telephones.

#### **Enhanced-Service Facility Providers**

The third and perhaps most important approach to the growth of the information-services market is the establishment of large, enhanced-services facility providers. Usually, these are either large service bureaus that have expertise with in-bound telemarketing, or long-distance service providers that have expertise in network-facilities operations. (*In-bound telemarketing* is the use of human or automated agents to answer

offer enhanced services from the network as long as the services use only *comparable and efficient* Interconnect (CEI) interfaces. That is, these interfaces to the basic network must be comparable to and as efficient as the interconnects sold to other enhanced-services providers.

incoming calls related to a business function, primarily sales and marketing of products or services.)

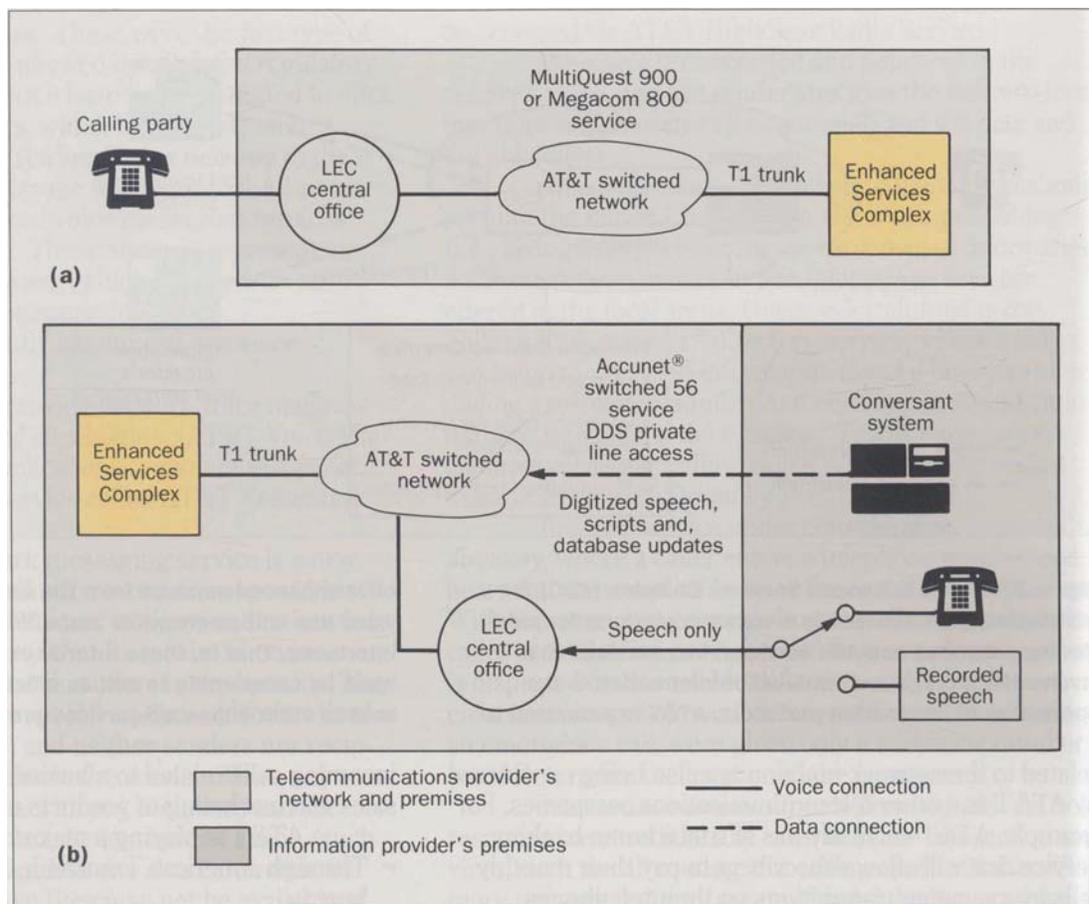
AT&T is playing a major role here in three areas:

- Through American Transtech, a fully owned service bureau
- Through Call Interactive, a joint venture with American Express to form a large, interactive-voice service bureau
- AT&T's recently announced InfoWorx<sup>SM</sup> family of interactive-voice services.

Although all three use the Conversant voice-information system as a key building block of their enhanced-services facility, only the InfoWorx interactive-voice services are network based.

The InfoWorx interactive-voice services include

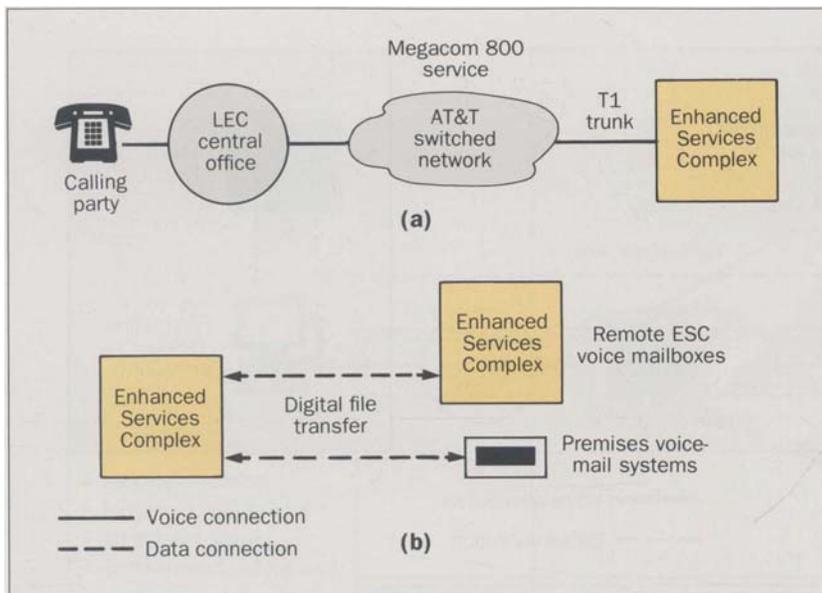
**Figure 3. Architecture for information and entertainment services. All information is stored in the ESC. (Top) Calls come into the ESC from either Megacom 800 or MultiQuest 900 service and terminate on a voice-response unit in the ESC. The VRU and caller interact through Touch-Tone signals and voice responses. (Bottom) The ESC allows frequent updates to the information provider's stored information. Updates may be digitized speech or entire applications (i.e., digitized speech, script, and databases) transmitted from an application development system.**



information, entertainment, and transaction applications that are directed primarily at the business-to-business or business-to-consumer markets. Through the InfoWorx services, AT&T is able to sell, develop, and operate a customized interactive-voice application associated with a business' 800- or 900-number service on a pay-by-minute basis. The InfoWorx interactive-voice services reduce barriers for businesses to enter the interactive-voice arena.

Along with AT&T Voice Mail, the InfoWorx

interactive-voice services have required that AT&T develop a network-based capability for handling such interactive-voice services. As its network platform for the delivery of the InfoWorx and AT&T Voice Mail interactive-voice services, AT&T has developed and deployed the Enhanced Services Complex (ESC). The remainder of this paper will describe the ESC and the technical architectures that allowed AT&T to introduce thousands of network lines of custom interactive-voice services in less than a year.



**Figure 4. ESC call architecture for voice-messaging services. (a) Message recording and retrieval, and (b) message forwarding stage for voice mail. AT&T Voice Mail service permits a caller to store a speech file in the ESC voice mailbox of the person being called. Later, the mailbox owner calls in to retrieve messages or create new ones to be forwarded to other mailbox owners. Message forwarding may involve data transmission of voice messages from one ESC site to another.**

### The AT&T Enhanced Services Complex

The Enhanced Services Complex (Figure 2) is the network-based platform for AT&T's InfoWorx and Voice Mail interactive-voice service offerings.

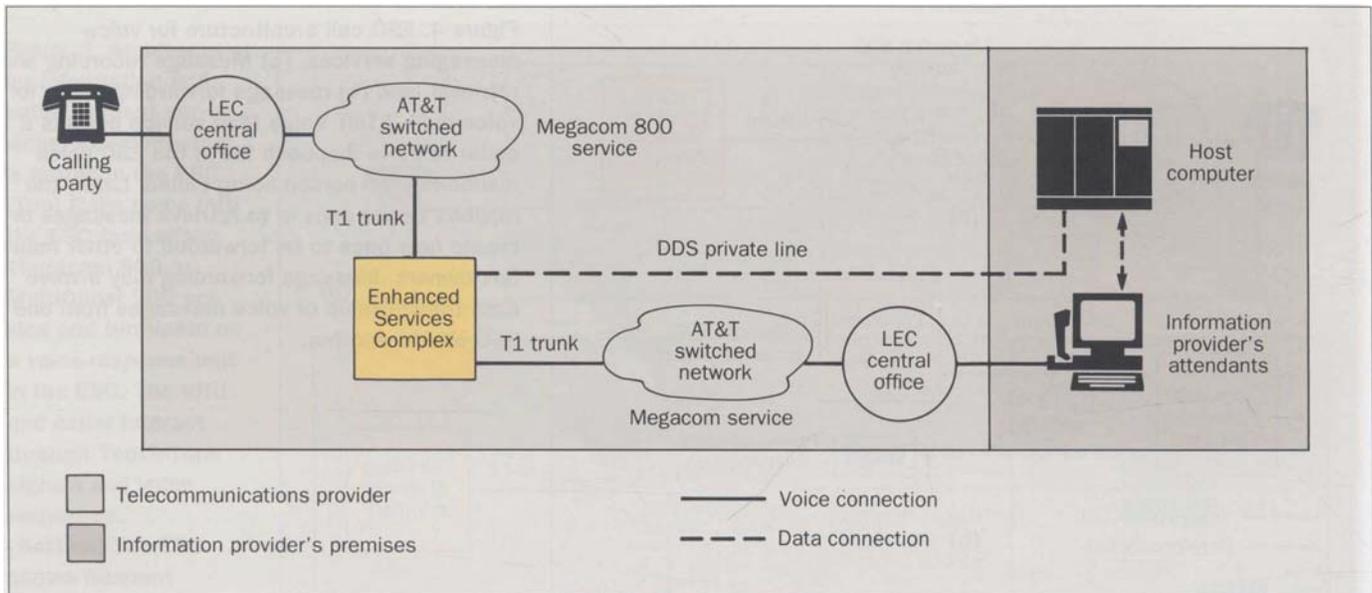
An ESC is a collection of switching, speech processing, and computer equipment located within a service node at the edge of the AT&T switched network. (*At the edge of the network* refers to the standard interfaces to our core network. These adjuncts answer calls that have passed from a caller to the LEC, through the AT&T switched network to its terminating "edge.") The ESCs are networked to provide enhanced voice and data services through standardized interfaces into public switched networks. These services will permit callers to interact with and manipulate various types of audio information stored in the ESC.

ESCs provide interactive-voice services via AT&T's tariffed nodal services including the Megacom<sup>®</sup> 800, MultiQuest<sup>®</sup>, and Megacom telecommunications services and the SDN. The ESCs are interconnected with

the AT&T switched network through tariffed interfaces, such as T1 facilities with ISDN primary-rate interfaces.<sup>5</sup> [ISDN is the Integrated Services Digital Network. In North America and Japan, primary-rate ISDN provides 23B+D, or twenty-three 64-kb/s (kilobits per second) circuit-switched channels (23B) and one 64-kb/s signaling channel (D). The B channels, also called *bearer* channels, carry the users' or customers' information (i.e., voice, data, and video signals). The D, or *data*, channel carries the network's control and signaling information. In Europe, the primary-rate arrangement is 30B+D.]

**ESC Applications.** A customer application consists of voice scripts, digitized speech files (i.e., the voice responses), application-specific databases, and customized software. To identify the application being called, AT&T's ESC uses a dialed-number identification service.

The call flows (i.e., the paths followed by calls) that involve the ESC depend on the type of interactive-voice service application being accessed (i.e., information, messaging, or transaction service applications).



**Figure 5. ESC call architecture for transaction services. Transaction applications access databases on the information provider's host computer. Because the ESC converts the call protocol (i.e., the Touch-Tone input and VRU voice responses) into the data protocol of the host computer, it can act as a pseudo-agent.**

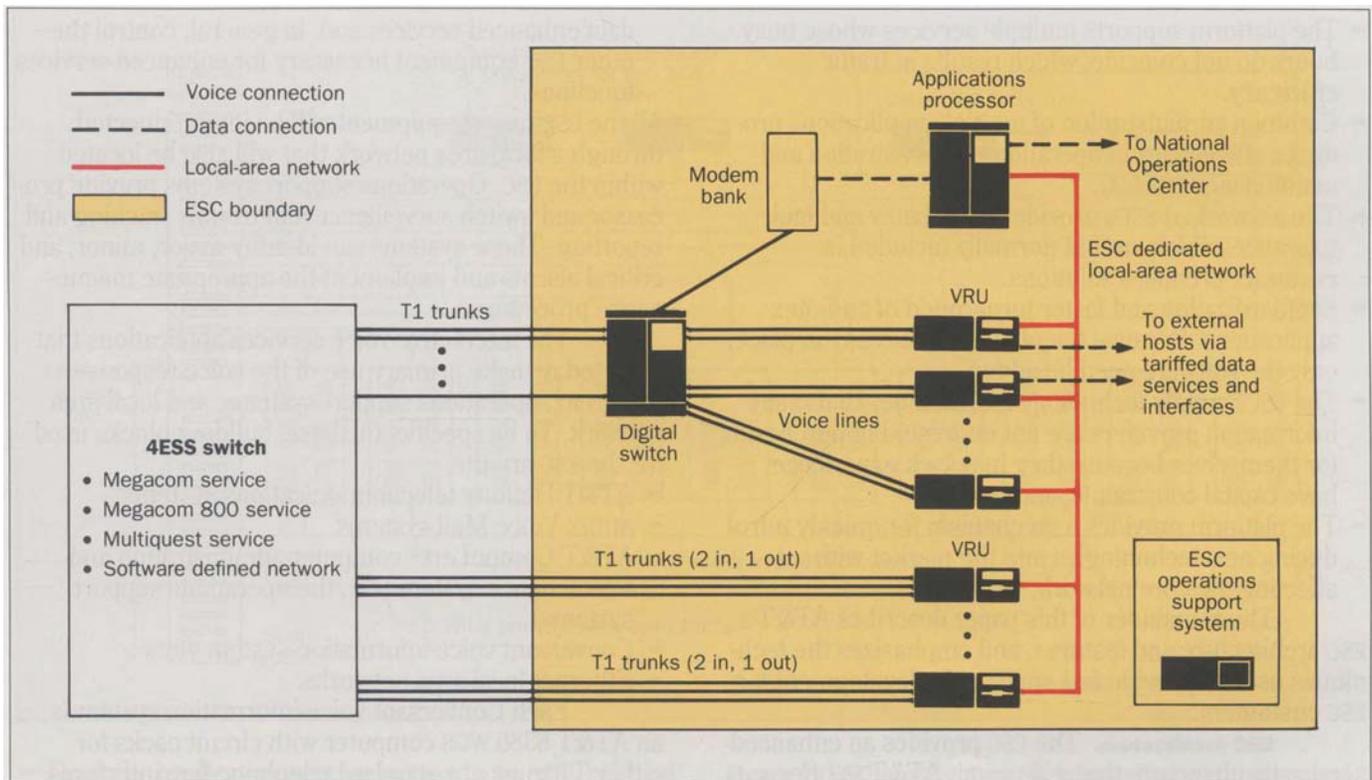
For an information distribution and entertainment service (Figure 3), calls come into the ESC from either Megacom 800 or MultiQuest 900 telecommunications service and terminate on the voice-response unit. The voice-response unit interacts with the caller through speech announcements, and accepts Touch-Tone signals from the caller's telephone.

All information is stored in the voice-response unit. For these services, particularly those that provide news, sports scores, and so on, it is critical that the ESC allow the information provider to make frequent updates to the stored information. These updates are either digitized speech or entire applications (digitized speech,

voice scripts, and the application's databases) transmitted from an application development system.

For example, each day, a horoscope service would require 12 new phrases of speech from the information provider. ESC technology allows the information provider to place a voice call into the ESC, record the new phrases, listen to them to assure their quality, and then have them distributed automatically to the Conversant voice-information systems that are answering the service's incoming calls.

AT&T Voice Mail is the ESC-based, voice-messaging service (Figure 4). This service involves an initial call, during which a caller stores a speech file in the ESC voice mailbox of the person being called. Later, the mailbox owner can call in and retrieve his or her messages, as well as create new messages to be forwarded to other mailbox owners. The message-forwarding capability may involve data transmission of the messages from the voice-response unit in one ESC site to the unit at another site.



Transaction applications (Figure 5) access databases that are external to the ESC. Because the data changes continually and usually is proprietary to the information provider, the databases must remain on the information provider's host computer. As a result, the ESC must convert the protocol of the application's Touch-Tone and voice-response interaction with the caller into the data protocol required by the information provider's host computer. This data protocol is usually the one used by the human-agent terminals that are connected to the information provider's host. The ESC's ability to act automatically as a *pseudo-agent* is extremely valuable to interactive-voice service customers.

Although the call flows differ greatly among the

**Figure 6. The ESC architecture will permit AT&T to offer a variety of voice and data services using equipment located in AT&T central offices. The digital switch routes incoming calls to and from the ESC's speech and data processing equipment. VRUs encode or decode speech outputs in response to a caller's Touch-Tone signals, and convert incoming messages to digitized speech for storage.**

three types of applications (i.e., information distribution, voice messaging, and transactions), there are many advantages to a common platform for provisioning enhanced network services. These advantages include:

- Reduced access and egress costs, compared to solutions that use premises-based equipment.

- The platform supports multiple services whose busy hours do not coincide, which results in traffic efficiency.
- Common administration of multiple applications produces efficiencies in operations administration and maintenance (OA&M).
- The networked ESCs provide redundancy and fault tolerance, which are not normally included in customer-premises solutions.
- Standardization and faster turnaround of audiotex applications. Because the platform is already in place, only the software need be added.
- The ESCs supply technology and facilities that many information providers are not interested in purchasing for themselves because they may lack experience, have capital constraints, and so on.
- The platform provides a mechanism for quickly introducing new technologies into the market without affecting the core network.

The remainder of this paper describes AT&T's ESC architecture and features, and emphasizes the techniques used to provide fast application development for ESC customers.

**ESC Architecture.** The ESC provides an enhanced-services architecture that will permit AT&T to offer a variety of voice and data enhanced services, using equipment located in AT&T central offices.

When collocated with the 4ESS™ toll switch, the ESC will consist of a variety of customer-premises-like equipment (Figure 6). For example, the ESC may contain:

- Digital switching equipment, such as a PBX, that will route incoming calls to and from the appropriate speech- or data-processing equipment within the ESC.
- Voice-response units to provide encoded or digitized speech outputs in response to Touch-Tone signals. The voice-response units will also record incoming messages, convert them to digitized speech, and store the speech.
- Application processors that will provide call-processing control of the digital switch, control the

data enhanced services and, in general, control the other ESC equipment necessary for enhanced-services functions.

All the ESC-based equipment will be interconnected through a local-area network that will also be located within the ESC. Operations support systems provide processor and switch surveillance and trouble tracking and reporting. These systems can identify major, minor, and critical alarms and implement the appropriate maintenance procedures.

The interactive-voice services applications that exist today make primary use of the voice-response units, PBX, operations support systems, and local-area network. To be specific, the basic building blocks used for the ESC are the:

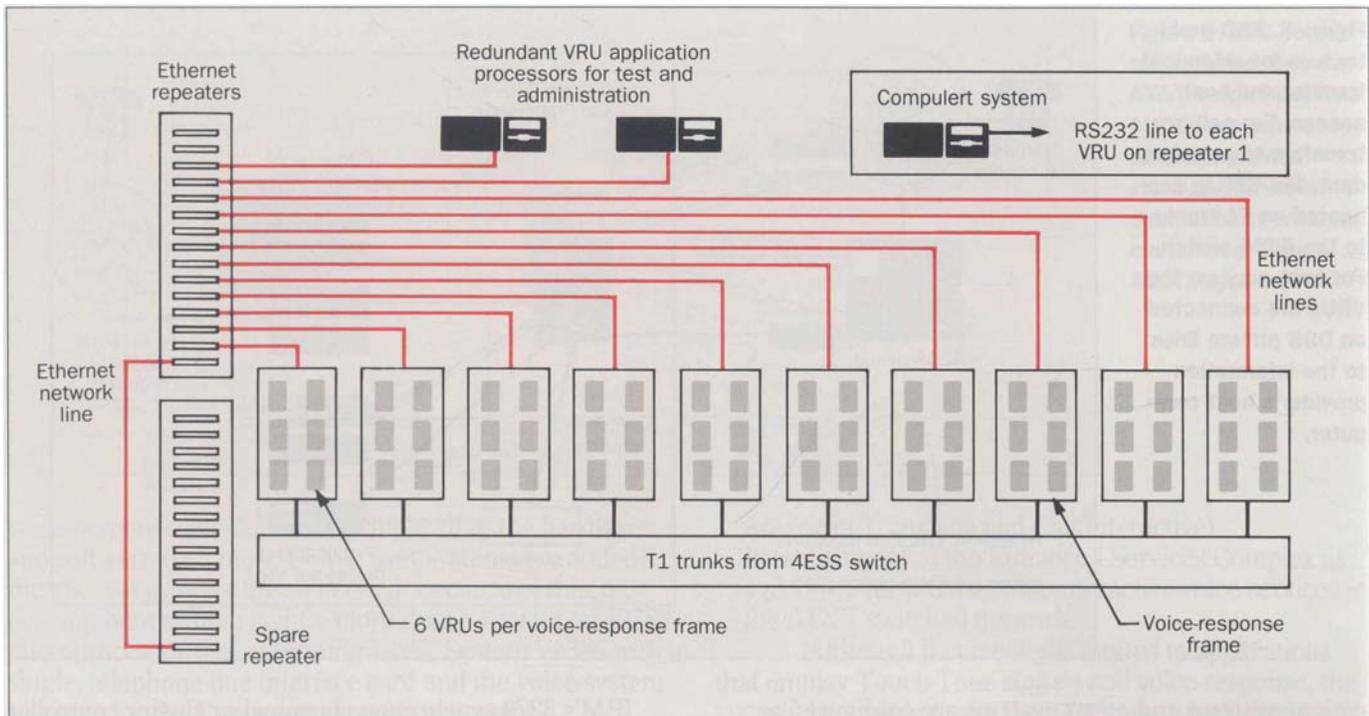
- AT&T Definity telecommunications systems
- AUDIX Voice Mail systems
- AT&T CompuLert® computer administration and maintenance system (i.e., the operations support systems)
- Conversant voice-information-system VRUS
- Ethernet local-area networks.

Each Conversant voice-information system is an AT&T 6386 WGS computer with circuit packs for either T1 trunks or standard telephone line interfaces. The AT&T 6386 WGS runs the UNIX System V/386 operating system. The Conversant voice-information systems are connected together on an Ethernet local-area network at 10 Mb/s (megabits per second) for collecting call-detail information, updating the applications, and generating reports.

Six voice-response units are installed in each voice-response frame (i.e., cabinet). In our current configuration, certain voice-response units provide the application-processor function of information collection, and the script and speech administration for the ESC.

The exact configuration of the ESC components depends on the answers to the following questions:

- Does the application require transfer of the call to a human agent or attendant?



- Does the application require access to an external host computer?
- Does the application require dedicated voice mailboxes?

For 800- and 900-number applications where all three answers are *no* (Figure 7), the voice-response frames are connected directly to the 4ESS toll switch with 12 T1 facilities (2 T1 trunks per voice-response unit). Because of the large number of voice-response frames, Ethernet repeaters are used, one repeater segment per frame. This reduces the effect of a local-area-network failure on any segment. These systems are currently in the ESC sites in Houston, Texas, and St. Louis, Missouri.

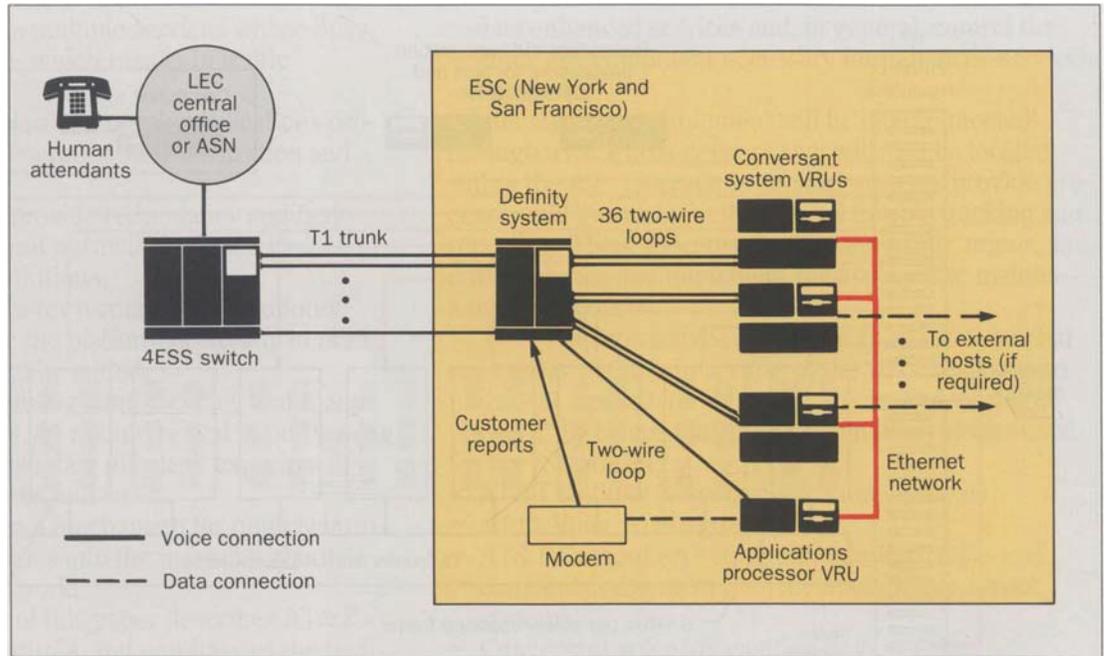
For applications that require call transfers to human agents, a Definity telecommunications system is usually used (Figure 8). Here, the voice-response units

**Figure 7. ESC architecture for information distribution and entertainment services. Voice-response frames are connected directly to the 4ESS toll switch using T1 facilities (two T1 trunks per VRU). The use of one Ethernet repeater segment per frame reduces the effect of a failure on a local-area network segment.**

connect with telephone lines to the PBX that, in turn, connects to the 4ESS switch over T1 facilities. For applications that require host access, the voice-response units are connected on DDS (digital-data service) private lines to the information provider's host computer. Usually, these applications also require call transfers to human agents. Most of these systems are in the ESC sites in New York, New York, and San Francisco, California.

For AT&T Voice Mail, the Definity telecommuni-

**Figure 8. ESC architecture for attendant transfer and host access. For call transfers to an attendant, the ESC is connected on T1 trunks to the 4ESS switch. For host access, the VRUs are connected on DDS private lines to the information provider's host computer.**



74

cations systems and AUDIX systems are configured as illustrated in Figure 9. The use of the AUDIX system as the voice-response unit provides voice store-and-forward compatibility with any AUDIX system on the information provider's premises. The ESC equipment for AT&T Voice Mail is also in New York and San Francisco.

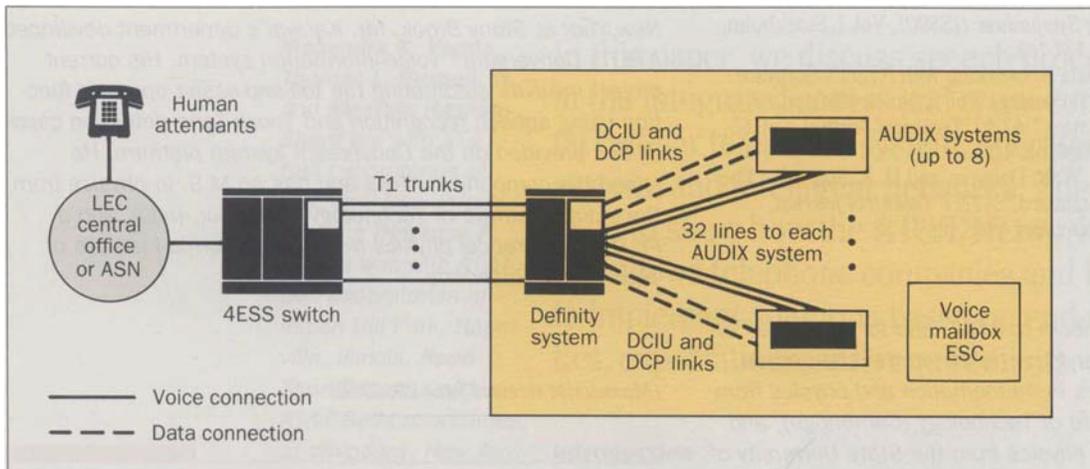
**ESC Features.** AT&T's ESC platform provides the following features for interactive-voice services:

- Voice encoding and decoding at 32 kb/s using adaptive-differential pulse-code modulation (ADPCM) technology.<sup>6</sup> The current ESC can support up to 65 hours of application speech. For voice-mail applications, 16-kb/s subband coding is used.
- Touch-Tone detection with the capability to interrupt voice announcements.
- Capabilities for interaction with an IBM host computer using the SNA™ 3270 network protocol. (SNA is a trademark of International Business Machines, Inc.

IBM's 3270 synchronous terminal or cluster controller system is an industry standard.)

- Hardware redundancy.
- Automated operations, administration, and maintenance.
- The option for customers to develop their own applications or have turnkey solutions provided. (See the discussion of the application development environment in the next section.)
- Large capacity to handle simultaneous calls.
- Call transfer to human agents.
- Standard and custom traffic reports.
- Voice-mail storage and forwarding.

**ESC Application Development Environment.** A major advantage of the ESC architecture just described is that the Conversant voice-information system is the service-providing equipment. Each Conversant voice-information system can provide up to 48 channels (2 T1 trunks) of



**Figure 9. ESC voice-mail architecture for AT&T Voice Mail. The AUDIX systems used as VRUs provide store-and-forward compatibility with any customer-premises AUDIX system.**

voice-response service and, by duplicating the hardware and software (i.e., more Conversant systems are added), the ESC can grow as big as needed. Because of this, a development system is little more than a PC with an 80386 microprocessor that is running UNIX System V/386 with a single, telephone line interface card and the voice-system software for the Conversant speech processor.

In addition, a high-level application generator called *ScriptBuilder* permits simple interactive-voice applications to be developed by someone with a minimal amount of programming experience or training. This makes it possible for information providers to develop and control their own applications on low-cost workstations, while allowing them to provide service on thousands of lines with redundancy from the ESC.

### Summary

The interactive-voice services market is now growing rapidly with a wide variety of information, entertainment, messaging, and transaction applications for both businesses and consumers. AT&T is playing a major role in the development of this market through:

- Equipment sales
- Participation in large service bureaus (currently,

American Transtech and Call Interactive)

- Establishment of the Enhanced Services Complex as a platform for provisioning interactive-voice services in the AT&T switched network.

Although it currently is limited to applications that employ Touch-Tone signals and voice response, the ESC will be evolving over the next year to include capabilities for speaker verification, speech recognition, and facsimile transmission. Through a common architecture and application structure, the ESC provides the technology to permit AT&T to provide a wide range of voice; voice and data; and eventually data, image, and video services—with the time to develop and deliver these services measured in months rather than years. It is an important step along the evolution path to the Information Age.

### References

1. A. F. Bulfer, "AT&T's Pay-per-View Television Trial," *AT&T Technical Journal*, Vol. 66, No. 3, May/June 1987, pp. 54-63.
2. R. J. Perdue, J. Leikness, and J. B. Sharp, "AT&T Voice Processing System Architectures," *AT&T Technical Journal*, Vol. 69, No. 5, September/October 1990, pp. 52-60.
3. M. Berger and P. G. Spring, "AT&T VoiceMark Messaging Service—A New Voice Store and Forward Service," *Proceedings of*

- 
- the International Switching Symposium (ISS90)*, Vol. I, Stockholm, Sweden, May 27, 1990, pp. 137-142.
4. L. C. Nicholson, "Virtual Data Networking with AT&T's Software Defined Network," *AT&T Technology*, Vol. 5, No. 3, 1990, pp. 38-44.
  5. R. T. Roca, "ISDN Architecture," *AT&T Technical Journal*, Vol. 65, No. 1, January/February 1986, pp. 4-17.
  6. N. Benvenuto, G. Bertocci, W. R. Daumer, and D. K. Sparrell, "The 32-kb/s ADPCM Coding Standard," *AT&T Technical Journal*, Vol. 65, No. 5, September/October 1986, pp. 12-22.

Biographies (continued)

*and financial issues of concern to AT&T, and for evaluation of long-distance consumer services. He joined the company in 1977 and has B.S. degrees in mathematics and physics from the Massachusetts Institute of Technology (Cambridge), and both an M.S. and Ph.D. in physics from the State University of*

*New York at Stony Brook. Mr. Kanwal's department developed AT&T's Conversant® voice-information system. His current project involves automating the toll-and-assist operator function using speech recognition and Touch-Tone detection capabilities provided on the Conversant system platform. He joined the company in 1983 and has an M.S. in physics from the Indian Institute of Technology at Kanpur, India, and a Ph.D. in theoretical physics from the California Institute of Technology, Pasadena.*

*(Manuscript received June 20, 1990)*

---