Telecommunications and Librarians

Karen Momenee

Applications of computer technology to libraries have come a long way in the past decade. We are no longer satisfied with batch processing and the attendant response delay. On-line interactive processing has become commonplace in reference, technical services and circulation. A practicing librarian wishing to find out more about this interactive communication process will find a wealth of material in the library and information science literature which focuses on the terminal end, or user side. There are clear explanations of the use of Boolean logic which provides the central operations for most search languages, the patron/searcher interaction in query negotiation, and the database vendors themselves provide extensive documentation on their retrieval language and file descriptions for various sources of information.

Perhaps less common, but still readily available, is information on the functions performed at the computer end of the communication chain. The inverted file no longer carries the mystique it once did. Library school courses often include sections on the underlying structure of the information contained on the computer's disk packs, and on methods of accessing this information by decoding search keys.

The curious practitioner will have a more difficult time in finding readable discussions of the middle ground, the rapidly growing field of telecommunications or more specifically, data communications, the process of sending information from terminal to computer, computer to computer, or, with the advent of intelligent terminals, from terminal to terminal.

A question was posed at a workshop recently, which highlights future possibilities for computer applications in libraries, but also emphasizes problems we may face in the 80's. A clearer understanding of data communications may provide the solutions to some of these problems.

The question was this: "In conducting an on-line reference search, is it possible to retrieve only those references to materials which you hold in your own collection?" A likely extension is to limit the references to those held in a specific set of libraries in your area, perhaps imposing a ranking function to order the references by library. This is certainly a reasonable enough request. The steps implied are:

1. Selecting the information from several databases;
2. Storing this information transiently;
3. Performing some calculations on the raw database information to come up with the final answer.

The picture I had in mind for step 1 above was access and selection of data from one of the commercial on-line vendors (Lockheed, SDC) and also access to an on-line union-list (locally or regionally generated, or through OCLC). In order to effect the automatic screening of the citations implied, some method would have to be developed for compensating for the different types of computers which might be accessed, each of which would have its own internal representation and structure for the desired information. The different computers might need different types of terminals for transmission and receipt of the information. The local processing done in step 3 above would require advance knowledge of the formats of the information to be received.

The major computer vendors (eg. IBM, DEC, Burroughs) have recognized these problems and along with national and international organizations concerned with communications standards (EIA, Electronics Industry Association, U.S.; CCITT, Consultative Committee on International Telegraphy and Telephony, International) are evolving methods for resolving the differences in the area of network architecture and distributed databases. The problem is divided into layers. Rules or protocols are developed for each layer. In general, network protocols are sets of rules which govern the flow of data in a network, within and between layers. The low levels are concerned with details on the physical or hardware side. The highest levels are more closely related to the logical applications of the user. Transmission between and with layers often involves "enveloping" or enclosing the actual information which is to be sent along the communications line and adding the address of sender and receiver, and a note on the length of the message being transmitted. Service programs at the appropriate level will read the contents of the envelope of concern, strip this information, and perhaps add yet another envelope as the message is sent on its way through the network.

Starting at the bottom in Figure 1, the lowest level protocol ensures that bits or pieces of coded data can be put onto the communications line and taken off again at the other end. Levels of interface specifications cover the number of wires in the connection and the mechanical arrangement and number of pins or plugs and sockets. The protocol describes how data are transferred across the interface in such detail that particular pins are designed as carrying specific messages at certain times. Two established protocols at this level are the EIA's RS-232 and CCITT's V24. Types of communications which would be exchanged at this interface include 'handshaking' or RTS—request to send (may I start dumping my bits on the line?)/CTS—clear to send (you may).

This is the type of protocol which governs the physical connection between a terminal and a modem. Modem is the abbreviated form of modulator/demodulator. A modem is needed if we are to use the public telephone lines for communications with a distant computer. Most telephone lines are set up to transfer voice grade analog signals. The signals generated at the terminal or at the distant computer, however, are digital, composed simply of combinations of zeroes and ones (bits) which represent alphanumeric characters when taken in bunches of 6, 7, or 8 (bytes). The modem generates an analog "carrier signal" and superimposes the bits to be transmitted on this wave. The modulation can be effected in several ways by changing frequency, amplitude, or phase. The bits are extracted from the carrier wave at the receiving end of the communication line by another modem.

Figure 1

Network Protocol Hierachy

Level	Name	Activities and Responsibilities
6	Applications	—Standardization of representation of actual library processes as referred to in applications programs
5	Presentation (Network/ User Interface	—Translation of information from/into required code, format for individual terminal or for the network itself
4	End-to-End Control (Network Services)	—Checks that all portions of message have been received and in the proper order.
3	Transport Protocol	—Path, Flow control; choice of path through network, assurance of smooth flow without overflow and loss of information at any node
2	Data Link Control	—Transit between two adjacent nodes including error detection and control
1	Physical Level	—Hardware electrical and mechanical interface

Figure 1 Examples of levels of protocol in a computer network. Terminology and division of layers differ from vendor to vendor. This diagram is a representation of a composite to give the flavor of the concept of hierarchial protocol levels.

This lowest level of protocol is not concerned with the integrity of the message. Error detection and correction is implemented at the second level, line (or link) control. At this level, protocols insure that a block of data gets from one end of a data link to another and that it gets there correctly. One method of automatic error detection is the cyclic redundancy check. In this method the same string of bits representing the message is interpreted as representing a number. The number is divided by a constant and the remainder of this division process is placed in the "envelope" surrounding the message. At the other end of the data link, the constant is also known, division is again applied, and the newly computed remainder is compared with the transmitted remainder. If they match, the data is considered to be correct and the message is processed or routed further along the network. If the remainders do not agree, bits have been lost or corrupted in transit, and the receiver may request re-transmission.

Let us consider the nature of this data link or line. It may be two-wire (half-duplex) admitting transmission in either direction, but only in one direction at a time. There is a delay in this type of line associated with the time it takes to turn the line around or prepare for response transmission in the opposite direction. Four-wire (full-duplex) lines avoid this delay by accommodating simultaneous two-way transmission, A communications line may be point-to-point or multi-point. In point-to-point communication, one line connects a single transmitter (eg. computer) with a single receiver (eg. terminal). In multi-point communication, one line connects a computer with two or more terminals. Not more than one terminal can transmit at the same time in the multi-point case, since the data from one terminal would collide with data from the terminal on the line and result in a garbled message at the receiving end. The flow of data in this case must be carefully controlled by the line control procedures established at the level two of the network protocol hierarchy. One method of line control is polling. The computer polls or questions the terminals one after the other, in turn, to determine whether the terminal has anything to send down the line. Each terminal is assigned a unique name. Higher priorities may be given to a terminal by calling its name more than once in the sequence, perhaps by returning to poll it after polling each of the other terminals. Thus, for a multi-point line with 5 terminals named A through E, instead of the sequence A-B-C-D-E, which would give each terminal equal priority, we might have the sequence A-B-A-C-A-D-A-E-A. This type of polling is sometimes called roll call polling.

The lines or links themselves may be leased (private) or switched (dial-up., public acces). With a leased private line, the connectin is always there, provided through a physical connection. Switching is familiar to us already through telephone calls. If we have a friend in California we call every week, the connection would appear to be the same each time, but the same physical lines would not always be used. There would be alternate lines available between two points, and there would be quite a few of these alternative paths along the total route from North Carolina to California. Minicomputers are often used to make the appropriate decisions about where to switch a message through the network nodes. Since the minicomputers are also capable of storing information, the 'store and forward' technique is often used in switching. This allows the network to act as a buffer, to hold mesages until line or links are open or terminals are ready to receive them. This storage capability also allows the network to compensate for varying speeds of transmission and receipt or display for the different types of devices connected to the network.

Dial-up public lines are provided through TELENET and TYMNET. These communications services are called VANS or Value-added networks. "Value" is added in the form of the automatic error correction discussed earlier, in the capability of re-routing messages when one or more link fails in the total network, and in the storage capacity used for buffering as described above. Since the thousands of terminals which can be connected to either network simulataneously will involve operator "think" time (when nothing is flowing down the line) as well as actual transmit time, it is worthwhile considering methods of mixing different messages to use the communication line as much as possible, rather than to allow one terminal to completely possess a channel during its connect time to the system. Packet-switching is a method used to increase line usage. Messages are divided up into small packets. The packets from several different terminals can be interleaved, and the line effectively shared by all of these terminals simultaneously. Because the packets which comprise one message may be sent along different paths, it is conceivable that they may reach their destination in the wrong order, or that some may be lost in transit. The use of packet-switching would thus require a level of protocol to ensure the integrity of the entire message.

In the library community, serious attention (in the form of committee work resulting in standards recommendations) has only been paid to the highest, applications level of data communications outlined in the NCLIS/NBS documents are only suggestions and do not carry the wide recognition and acceptance of the other protocols discussed above. The first step toward standard and protocol development must surely be an informed library community, and regrettably we have only just begun to contemplate that first step.

Karen Momenee spoke at the May 1980 conference sponsored by the UNC-CH Librarian's Association. She distilled her remarks at the request of the editor, to point to a direction for computer technology in the 80s. Recently Ms. Momenee was employed by IBM, Raleigh, as an associate programmer.

SELECTED BIBLIOGRAPHY ON DATA COMMUNICATIONS

I. From the library community:

Avram, H. D. and D. C. Hartmann. "Objectives and accomplishments of the Network Technical Architecture Group," *Program*, 13(1):4, 1979.

NCLIS/NBS Task Force on Computer Network Protocols. *A Communications Network Protocol for Library and Information Science Applications*. Washington, D. C., 1977.

II. Periodicals

a. General Automation

Datamation (monthly) Technical Publishing Company

—Free to "qualified individuals ... employed by companies involved with automatic information handling equipment"

—Well worthwhile for the advertising and "News in Perspective"

—Each issue has a theme—July '78 and March '79 were data communications

b. Telecommunications

Communications News (monthly) Harcourt Brace Jovanovich

—"All that's new and newsworthy in voice, video and data communications"

—May, 1980 issue—Special Report: Data Communications

IEEE Transactions on Communications

—Vol. Com-28 No. 4, April 1980—An entire issue on Network Architectures and Protocols

III. From the computer perspective:

Davies, D. W. and Barber, D. L. A. et al. *Computer Networks and Their Protocols*, Wiley, 1979. (Very complete and thorough with a British point of view, includes a small section on Teletext and Viewdata)

Doll, D. R. *Data Communications: Facilities, Networks and Systems Design*, Wiley, 1978. (A thorough discussion from the American point of view, he also writes for *Datamation*)

Housley, T. *Data Communications and Teleprocessing Systems*, Prentice-Hall, 1979. (Starts out slowly at a basic level, but gives thorough coverage to the major concepts)

Martin, James (If you have not yet discovered the prolific and very readable James Martin, here are two recent titles to start with:

Telecommunications and the Computer. 2d ed. Prentice-Hall, 1976.

The Wired Society. Prentice-Hall, 1978.

Swartz, M. *Computer-Communication Network Design and Analysis*. Prentice-Hall, 1977. (This rapidly becomes mathematical, but the 2d chapter gives a good overview of the design of several functioning networks, including TYMNET)