[From: *Tools for the trade: Translating and the Computer 5*. Proceedings of a conference ...
10-11 November 1983, ed. Veronica Lawson (London: Aslib, 1985)]

# Development of computers and communications: the role of the National Physical Laboratory

*Dr Wyn L Price*

*Principal Scientific Officer, National Physical Laboratory, Teddington, UK*

*The paper traces the computer research programme at NPL, from the early development of computers, through various applications (including natural language processing), to computer networking. It then surveys some of the computer network services now generally available.*

## INTRODUCTION

The National Physical Laboratory (NPL) is a research establishment of the Department of Trade and Industry. The Laboratory was founded in 1900, due largely to the influence of the British Association for the Advancement of Science and the Royal Society. The original remit of the Laboratory laid strong emphasis on the creation and maintenance of measurement standards, calibration and related physical research; activity in this area remains a major part of the work of NPL today, but the Laboratory has grown and its research has diversified well beyond its first objectives. A detailed history of the Laboratory by Pyatt (1) has been published recently.

Since 1945 the Laboratory has played a leading part in the development of computers and communication involving computers. Work began in 1946 on an 'automatic computing engine', adopting a name originally used by Charles Babbage, 'father' of computing, which led to the building of the Pilot ACE computer. This began working in May 1950 and was one of the world's first computers; at that time it was also one of the fastest. Pilot ACE gave sterling service to the Mathematics Division at NPL for a number of years; its work included studies of road traffic control and aircraft design. When it was dismantled at the end of its service it was transferred for display at the Science Museum.

Pilot ACE was succeeded by DEUCE, an engineered version of the same basic design by the English Electric Company. In turn this was itself replaced by the ACE computer, designed and built at NPL, a much more powerful machine which came into service in 1958 and was heavily used until it was closed down in 1967.

The 'full-scale' ACE was the last of the series of machines designed and built at NPL. Today the Laboratory has a large number of computers, ranging from micro-processors to big 'number crunchers', which play a vital part in the general research work; these come from a wide range of manufacturers. The teams that built the early NPL machines, together with new staff joining over the years, turned their attention to applications of computer systems; these applications were generally 'non-mathematical', such as language translation and pattern recognition. At the same time the mathematical work expanded in the Mathematics Division.

## NATURAL LANGUAGE PROCESSING AT NPL

The machine translation research programme began in 1959 and aimed at translating Russian scientific texts (particularly in the fields of electrical engineering, mathematics and physics) into English, using the 'full-scale' ACE computer. The standard of translation aimed at was one that would be useful to a professional worker in the particular field; translations of literary quality were not sought. A dictionary of about 20,000 entries was created, first punched on cards and then written on magnetic tape. This dictionary was organised by Russian 'stem' forms, with descriptions of the inflections possible with each stem. Allowance was made for the systematic stem changes which occur in Russian. Each entry contained at least one English equivalent, with additional equivalents where alternatives were possible.

The translation process commenced with a scan of the Russian text designed to split stems from suffixes. This enabled the words to be looked up in the tape dictionary. Also at this stage Russian word groups forming idiomatic phrases were identified and the corresponding dictionary entries found. Syntactic analysis programs were written which allowed much of the structure of the Russian texts to be discovered. A complete syntactic analysis was not aimed for. Other programs took the partially analysed text and created the corresponding English structures. Into these structures the English equivalent forms were fitted. The result was punched on cards for output from the computer.

Finally the cards were tabulated on a card-operated type-writer. Where more than one English equivalent was possible, these were output as alternatives to be selected by the reader. Where no dictionary match was found, the Russian word was transliterated into 'English'; because of the large scale 'borrowing' of English words to make Russian technical terms, this transliteration often gave a perfectly readable 'word'.

At the end of the machine translation research project, independent assessment of the product quality was sought. Scientists were invited to send in Russian texts for processing by the NPL translation system. The results were then sent back to the originators for evaluation against a nine-point scale which ranged from 'useless' to 'fully adequate'. Altogether thirty-four texts were subjected to this process and the average response corresponded to the description 'mostly very good - a few sentences obscure, so that something may be lost, but normally clear enough'. A description of the NPL machine translation project may be found in the paper by McDaniel et al. (2)

The machine translation project was followed by another project in the field of natural language processing. This was intended to take machine shorthand and convert this into readable English. The project was based on the Palantype system of mechanical shorthand which employs a machine with a 29-key keyboard; the keys are arranged in groups of left consonants, vowels and right consonants and the recording system follows fairly closely the syllabic structure of words, though some shorthand 'chords' cover more than one syllable. The normal output of the machine is a strip of printed paper from which a transcription is made by a human operator, a time-consuming operation which the NPL project aimed to obviate.

The Palantype machine was adapted at NPL to permit data input to the KDF9 computer. The next task was to create a Palantype-to-English dictionary. A suitable word list was obtained and this was Palantyped and recorded on punched paper tape. From this the machine dictionary was built, taking into account the syllabic structure of the Palantype system. As with the Russian translation system, it was sometimes necessary to include more than one English equivalent; these arose because of homophones.

In action the transcription system accepted a stream of Palantype chords and looked these up in the dictionary to find the English equivalents. Because word boundaries were rarely explicit in the incoming stream, the process of dictionary lookup had the additional task of trying to identify word boundaries; this was done on the principle of 'longest-match', which usually (but not always) gave the right result. The English output from the transcription system was printed on a fast character printer. Where dictionary lookup had failed to find a match, the process took the Palantype chords and transliterated these, often giving a readable equivalent, just as was done with the Russian translation.

Given an error-free input, the transcription system performed very well, but fast operation of the keyboard at normal recording speeds resulted in a large number of input errors with which the system could not cope adequately. Before the project ended techniques were being developed which introduced a degree of error correction into the input text, leading to better-quality output. An account of the NPL Palantype transcription project may be found in Price. (3)

Though the NPL project ceased in 1970 after four years work, it is pleasant to be able to report that others have since built on this work (4, 5, 6), leading to useful applications in subtitling television programmes and in aids for the deaf. These later developments have made use of microprocessor systems, which were not available at the time of the NPL project.

Both the Russian translation and the Palantype transcription systems depended on very large computers, and this diminished their practical application. The Russian translation system was essentially a batch process, whilst the Palantype system could operate either online or in batch mode. Batch processing was the norm for the vast majority of early computer applications, with jobs being submitted in the form of packs of program and data cards (or tape) to machine operators. Computers usually handled only one job at a time.

As computer technology developed, so we saw the introduction of time-sharing, in which more than one task at a time was handled by a machine, each task getting its turn under the control of the computer operating system; thus the main processor of the machine might be occupied with one job, whilst a backing-store data transfer was going on for another job. With the possibility of running multiple tasks the need for presenting jobs in batch form became less insistent, and the computer users were able to prepare programs and submit jobs from terminals without directly involving the computer operators; results were often delivered to the user terminals. This development, with users becoming increasingly remote from the computers, led to exciting innovations in communication technology.

The level of innovation increased enormously when communication systems were required to allow computers to communicate with each other and not just with remote terminals. The era of distributed processing has meant that some of the more complex computational tasks are now performed using computing capability located at more than one site, with co-operating computers and databases.

## THE DEVELOPMENT OF COMPUTER COMMUNICATION

A possible way to connect a number of user terminals to a computer mainframe is to provide each with a separate electrical connection, or line, and bring these lines into the computer centre, where a front-end processor has the task of handling them. Indeed, this is how the first terminal systems were connected. For a few terminals installed within a limited site the method is feasible. However, for a large number of terminals installed over a wide area the quantity and cost of dedicated hardware becomes frightening.

It is therefore necessary to look for alternatives which will allow some sharing of systems hardware between terminals. One very familiar system in which many people share communication hardware is the public switched telephone network (or PSTN), which has been with us for many decades. For most of its lifetime this system has supported only voice communication, but within the last twenty-five years or so it has been increasingly used for data also. For this purpose data is converted into sound pulses at the transmitter and back into data waveforms at the receiver; the device which carries out this function is known as a 'modem' or modulator-demodulator.

Unfortunately there are severe limitations on the data capacity of the traditional telephone network, which have restricted its usefulness in computer communications. Another important limitation arises from the time taken to set up calls in the PSTN. Where computers (or terminals) need to communicate sporadically with each other, using bursts of data, it is very inefficient to set up a fresh call every time a data burst must flow; the exchange and line allocations are held for the full duration of a PSTN call and will be underused for bursty data applications. A typical application of this kind is the modern automatic teller machine, seen outside banks, or the point-of-sale terminal which will soon be with us. On the other hand the switched system may well be able to handle large quanitities of data in a continuous flow, albeit at fairly low data rates. There is clearly a need for an alternative communication mode.

Such an alternative system is now with us in a fully-developed form and its inception is largely due to another initiative from NPL. In 1965 a group under Davies developed the concept of 'packet-switching' as an alternative to the traditional 'circuit-switching' of the PSTN. The PSTN consists of a large number of exchanges joined by a network of lines. Dialling a call in a circuit-switched system causes a path to be found from source to destination user, with each exchange along the path taking its part in establishing the chosen route; the path is then reserved for the duration of the call and the data flows along the reserved path, line and exchange hardware being dedicated to this task.

In contrast, a packet-switched system requires that data be segmented into units called 'packets'; each packet carries a header field in which are recorded such parameters as source and destination designation, packet number, etc. The packet-switched network consists of a number of nodes (small computers) joined by a network of lines. As a packet arrives at each node on its journey from source to destination, so the packet header is inspected and the node decides on which output line the packet shall leave. When the packet arrives at the destination node it is delivered to the user. Packet switching may take many forms; routing may be adaptive, taking account of changes in network topology or in loading levels, or it may be fixed, in which case packets between particular sources and destinations always take the same paths; flow control of packets may depend on end-to-end exchanges of control signals, or it may be based on node-to-node control or some combination of these procedures. Error detection and recovery is an important feature of these systems; again this may be based on end-to-end or node-to-node procedures.

NPL developed the concept in the late 1960s and soon had its own on-site packet-switched network in operation. This provided access within the Laboratory to services and terminals; at first it was run on an experimental basis, but soon it came to be relied upon as an essential feature of the computing systems of the Laboratory. In addition to giving access to the central computing facilities at the Laboratory, the network gives access to other systems such as the word processing facilities, Scrapbook and Edit. The network at NPL is still in operation, the present system being a direct development from the first. We see thus that NPL had its own packet-switched local area network as much as fifteen years ago. A description of the NPL network may be found in Davies and Barber. (7)

At the same time packet-switching developments began in the United States with the network of the Advanced Research Projects Agency (ARPA) of the Department of Defense. This rapidly expanded from four nodes in California to sixty-four nodes covering the whole of the United States and beyond. Since that time many more networks have come into existence and carry an immense worldwide load of data traffic. Commerce and industry have come to rely very heavily on network systems.

During the first years of development all that the different networks had in common were the basic concepts.

Any intercommunication between networks had to be achieved by means of specially designed 'gateway' nodes. This was largely because there was no commonality of procedures or of data formats. At the gateways the protocols of one network were translated into those of the next. This was clearly a very undesirable state of affairs, because ad hoc gateways were required between all network pairs needing to communicate.

The remedy for incompatible networks is to establish a set of standards to be followed by all network designers. Standards for networking began to emerge quite early. One of the first was British Standard 4421, which defined a 'digital input-output interface for use in data collection systems'. The most significant source of recommendations in the field of data communications has been the International Consultative Committee for Telephones and Telegraphs (CCITT), part of the International Telecommunication Union (UIT), itself an agency of the United Nations; the CCITT V series of recommendations is concerned with data communication via the telephone system; the CCITT X series of recommendations is concerned with the new data networks. Though going under the name of 'recommendations', the CCITT documents have almost the same status as formal standards; they are followed closely by the national PTT (Post, Telegraph and Telephone) organisations.

The emphasis of the CCITT recommendations is towards the hardware of the networks, so that one finds recommendations for signal levels on lines or for methods of error correction on point-to-point connections; one does not find many recommendations for protocols for allowing meaningful and efficient communication between one application program and another.

In order to achieve effective communication between computers or between terminals and computers, it is necessary to define a hierarchy of protocols. For example, there will be a protocol for data exchange between a terminal and the network node to which it is connected; there will be another protocol to establish reliable communication between source and destination hardware. Recently the International Standards Organisation through its Technical Committee 97, responsible for data communications, has created an architectural model, known as Open Systems Interconnection (OSI), whose function it is to provide an effective framework into which the individual standards can be slotted. The object of OSI is to allow inter-networking in the broadest sense. The architectural model has seven layers, ranging from the lowest (Physical) layer to the highest (Application). The Physical layer includes the mechanical and electrical characteristics of point-to-point connections, whilst the Application layer provides for interfacing to user application programs.

Definition of standards at the various OSI layers is proceeding apace; broadly speaking the progress is bottom upwards, with many of the CCITT recommendations forming the basis of international standards.

## NETWORK SERVICES AND APPLICATIONS

Though we may not be aware of the fact, many of us already use communication network facilities. Bank automatic teller machines rely upon network connections for reference to customers' accounts. Soon we shall see point-of-sale terminals appearing in the shops, so that 'plastic money' starts replacing cash for many transactions. These too will rely upon data networks to make the money transfers that are consequential upon purchases.

The banks already rely on data networks for internal banking transactions. Soon we shall see the CHAPS network (Clearing Houses Automatic Payments System) coming into action for very large inter-bank money transfers within the City of London, whilst the SWIFT network has carried international inter-bank transactions for a number of years.

Industry too relies heavily upon networking for communicating between various headquarters, offices and production plants.

Office automation will be greatly facilitated by the Teletex service, offered by British Telecom and a number of other PTTs. This is a document handling system, defined by the CCITT, which allows sophisticated document formatting and transfer between users. Teletex, not to be confused with teletext, could be described roughly as a 'super-telex' service; indeed it is intended to work alongside the telex system and to inter-work with it.

To the domestic user the teletext service (Ceefax and Oracle), available without additional charge on television, represents a primitive data information system; interfaces are now available between teletext and personal computers, so that programs and data may be loaded direct. At a more sophisticated level we have the Prestel system, also displayed on TV sets, but requiring a telephone connection; Prestel allows the user to operate in an interactive mode, so that, for example, goods may be ordered from advertisers. Already home banking is a possibility to a limited extent via Prestel, and may be expected to become more widely available in a short time.

There is already considerable international interest in the concept of an Integrated Services Digital Network. British Telecom plans to introduce such a service in late 1983. (8) This network will provide customers with a variety of new services and facilities, many of which are made possible only by the increased bandwidth provided by a wholly digital connection.

Some of the facilities and services we have mentioned use the older established media such as the telephone, but most of them would not be possible were it not for the development of the digital computer. It is truly astonishing that all this has come about within the thirty-eight years that have elapsed since the end of World War II. The pioneers of computing, imaginative people though they were, can have had little conception of the developments that would result from their work.

REFERENCES

(1) PYATT, E. The National Physical Laboratory - a history. Bristol: Adam Hilger Ltd, 1983.

(2) McDANIEL, J., PRICE, W.L., SZANSER, A.J.M. and YATES, D.M. An evaluation of the usefulness of machine translations produced at the National Physical Laboratory, with a summary of the translation methods. In: Proceedings of the 2nd International Conference on Automatic Translation, Grenoble, August 1967 (Session I, paper 2).

(3) PRICE, W.L. Palantype transcription by computer - a final report. National Physical Report, COM45, February 1971.

(4) DOWNTON, A.C. and NEWELL, A.E. An assessment of Palantype transcription as an aid for the deaf. International Journal of Man-machine Studies, 11, 1979, 667-80.

(5) HAWKINS, W.R. and ROBINSON, R.N. The development and use of an electric keyboard for television subtitling by Palantype. International Journal of Man-machine Studies, 11, 1979, 701-10.

(6) BOOTH, A.W. and BARNDEN, M.S. Voice input to English text output. International Journal of Man-machine Studies, 11, 1979, 681-91.

(7) DA VIES, D.W. and BARBER, D.L.A. Communication networks for computers. London: John Wiley & Sons, 1973.

(8) HARDY, J.H.M. and HOPPITT, C.E. Access to the British Telecom ISDN. In: Proceedings of the 6th International Conference on Computer Communication, London, September 1982, pp. 43-8.

AUTHOR

Dr Wyn L. Price, Principal Scientific Officer, National Physical Laboratory, Teddington, Middlesex TW11 0LW, UK.