XXIV. COGNITIVE INFORMATION PROCESSING

Academic and Research Staff

Prof. Jonathan Allen
Prof. Barry A. Blesser
Prof. Murray Eden
Prof. Francis F. Lee
Prof. William F. Schreiber
Prof. Hoo-min D. Toong
Prof. Donald E. Troxel

Prof. Ian T. Young
Dr. Christine E. Choppy
Dr. Takeo Eguchi
Dr. Denis J. Kfoury
Dr. Tadashi Mimaki
Dr. Douglas O'Shaughnessy

Dr. Donald G. Ross
Dr. Stephanie E. Sher
Dr. Robert J. Shillman
Caleb J. Drake
M. Sharon Hunnicutt
Robert A. Piankian
Kenneth P. Wacks

Graduate Students

Mark M. Abrams
Russell R. Atkinson
Wayne D. Baron
Robert P. Bishop
Ronald E. Boucher
Robert R. Buckley
Geoffrey J. Bunza
Nancy J. Burzinski
Robert E. Bruccoleri
John C. Carter
Yao-Ming Chao
Charles H. Cox
Paul J. Curlander
Mark Dijon
Peter V. Dolan

James R. Ellis, Jr.
Irvin S. Englander
Edward M. Garber
Richard S. Goldhor
Samuel M. Goldwasser
Robert E. Green
Randolph B. Haagens
Monson H. Hayes
Gerald L. Hoover
Su Ling Hsu
Malik M. A. Khan
Gary E. Kopec
Theodore T. Kuklinski
Peter LaMaster
Robert C. Laurenson

Donald S. Levinstone
Charles W. Lynn
Naveed A. Malik
Glen S. Miranker
James F. O'Toole
Leonard Picard
John N. Ratzel
James R. Seeley
Daniel G. Shapiro
Jorge Sianez-Guiterrez
Jonathan M. Teich
Victor Tom
Gary V. Vacon
Lindsay A. Weaver
Peter Yanover

1. NATURAL LANGUAGE PROCESSING

Jonathan Allen

The overall objective of this project continues to be the development of both a comprehensive model for converting English text to speech and a practical implementation of these techniques in special-purpose hardware.

The entire model is a set of algorithms (currently coded in BCPL) with well-defined input-output interfaces. Each of these is devoted to a particular aspect of text analysis or speech synthesis, including morphemic analysis, parts-of-speech computation, parsing, letter-to-sound conversion, lexical stress determination, pitch and timing computations, and, finally, phonemic synthesis of speech by rule. Our efforts are now focused on the last two parts of the system. We have implemented preliminary versions of timing and pitch algorithms, but these areas continue to require further research. There are a large number of determinants of pitch and timing, based on all aspects of the initial text and derived phonetic string. Thus it is a difficult task to determine an algorithmic specification for pitch and timing, but even more difficult to determine the perceptual necessity of these correlates. Furthermore, it has become clear that certain

linguistic features, such as stress, have multiple correlates that trade among each other in a way which is not currently understood. This complex relationship is probably an important factor in determining the naturalness of synthetic speech. Since the level of intelligibility of our synthetic speech is already high, the provision of naturalness for satisfactory long-term use becomes important. For these reasons, a major effort is being made to determine the relation between linguistic structures and these prosodic correlates, and also to express these relations in a well-structured algorithm appropriate to the needs of practical synthesis of speech by rule.

As a research vehicle, the model for text-to-speech conversion that we have developed is flexible, easy to understand and modify, and relatively machine-independent. On a conventional computer, however, it will not work in real time, and such an implementation would be too costly and bulky for most applications. For this reason, attention has now been directed to the development of custom-integrated circuits for practical implementations. Starting in 1978, we shall undertake the design of the entire phonemic synthesis part of the system in custom MOS technology. This project will require the design of custom architectures for these procedures, as well as complete logical design and layout. At present, the terminal analog synthesizer used to generate the output speech waveform requires approximately 150 dual-in-line circuit packages. Our expectation is that this task can be accomplished using 6 or fewer custom packages, with corresponding savings in space, power, and cost. We are also examining the possibility of custom devices for conversion of the phonetic transcription to parameters for the synthesizer circuit. The computational load for this task has been assessed, and an initial architecture is being devised.

The eventual goal of our work is to provide practical means for the generation of highly intelligible and natural-sounding speech from text in a fundamental way based on sound results from linguistics and speech science. This aim has led to parallel efforts in both basic research and custom implementation techniques. Our belief is that this strategy will lead to high-quality, practical implementations within the next few years, which will be attractive for a broad range of computer-based applications.

2. CONSTRUCTION AND GRAPHIC DISPLAY OF GUJARATI TEXT

Graphic Arts Research Foundation (Grant)

Francis F. Lee, Firoz Dosani

a. Introduction

This report describes the work done to design and implement a set of computer algorithms that would permit the construction and graphic display of Gujarati text on a relatively simple keyboard. By typing keys on a keyboard such as an ordinary teletype

connected to a computer system, one would produce as output, text (words) in the Gujarati language. One approach to producing an integrated system has actually been implemented. This includes techniques for constructing, sorting, and selectively retrieving for display the characters of Gujarati. The procedure used is described below.

b. The Gujarati Writing Form

Gujarati is an Indian language whose writing form and structure is very typical of most other Indian languages such as Hindi. The basic unit of the Gujarati writing form is the "aksara" or open syllable. Also included as an essential and inherent part of the language are vowel and diphthong markers. These generally occur with every consonant to denote its pronunciation. Gujarati is further enhanced by the great number of consonant conjuncts.

c. Approach

One very significant idea in the approach used was to take what is typically thought of as the basic structure or element of the writing form — the aksara or open syllable — and break it into finer units. Thus the first task was to determine these finer units or basic building blocks (graphemes) from which the aksaras of the language are constructed, and define combining rules to form the aksaras from these graphemes. One of the reasons for this breaking down is the consonant conjuncts feature of the language. Most of the consonant clusters are constructed by taking a part of the first consonant and combining it with the second consonant in the cluster. This part of the first consonant would be one of the graphemes or finer units used to represent that consonant. Also, many of the syllables have a common subunit, so that the use of basic building blocks to construct the syllables would result in considerable savings of the time and storage space required to define all the clusters graphically.

The next step was to write computer programs to construct and display the basic building blocks for the aksaras, to combine these to form the actual aksaras, and to construct and display the vowels and diphthongs. This was done in PL/1 on Multics, primarily because Multics supports a good graphics system that is highly terminal-independent. In writing programs to define the graphemes, an important design criterion was that the aesthetic quality of the output text would have to be of a sufficiently high standard to be acceptable to the Asian peoples, most of whom have a highly developed taste for good calligraphy. This consideration led, after some experimentation, to the decision to define the graphic elements using a dot matrix representation rather than a vector matrix representation. Further experimentation showed that an 8 * 8 dot matrix representation did not produce the fineness that was desirable to distinguish clearly between any two similar characters in the language, and so a 16 * 16 dot matrix was preferable. The extra cost of storing all the graphemes (up to four times as much

memory) seems to be worth the resulting elegance and naturalness of the characters. Also, since some aksaras are much wider than others (even by more than a factor of two), rather than using a 16 * 16 square matrix representation, it was decided to keep the width of the aksaras variable, going from as little as 8 up to a maximum of 32 units. In addition, there would be a variable but maximum of 12 units above, and 8 units beneath an aksara, for the vowel and diphthong markers.

The third and final major step involved determining and defining rules by which text (words) is constructed from the aksaras, the consonant clusters, and the vowel signs. This includes a fairly complex computer program for the selection and correct positioning of the vowels. Although most of the aksaras are internally stored in two distinct parts, externally, on the teletype, only one key is used to denote the full aksara; so that typing a key denoting a particular syllable will cause both parts of the syllable to be output and correctly positioned, whereas typing a key for a consonant cluster may cause only one of the internal parts of the first consonant to be output. This was done because there are far too few keys on an ordinary teletype to allow the use of separate keys to denote the different parts of an aksara. Separate keys are used to denote the vowel and diphthong markers. These markers have to be positioned correctly by the program, to the right of, to the left of, above, beneath, or partly overlapping an aksara or a consonant adjunct. Other special features of the writing form are also taken care of.

d. Results and Conclusion

The user-visible program will take in a string of characters, and output in Gujarati the aksaras, the consonant conjuncts, and the vowel and diphthong markers that the character string denotes. Moreover, spacing of the output text — for instance, the spacing

ASSS C4 k4 ⌄2DST1!A6 g3jr5t1Ls5h2t2y ⌄SMc1 t4n6 ABZy5s krS4.

આશા છે કે વિદ્યાર્થીઓ ગુજરાતી

સાહિત્ય વાંચી તેનો અભ્યાસ કરશે

Fig. XXIV-1. Translation: I hope that students will read Gujarati
literature and study it.

between lines, between words and between letters — as well as the size of the output characters, are parameters that the user can specify and change from one invocation of the program to another. A sample two-line output text is shown in Fig. XXIV-1. Further

work in this area could lead to a development of a computerized type-setting system to be used by the newspaper and other printing industries of the third-world countries, for the printing of Gujarati and of other similar languages.

3. DATA PROCESSING FOR THE GRAPHIC ARTS

Providence Gravure, Inc. (Grant)

William F. Schreiber, Donald E. Troxel, Leonard Picard, Peter V. Dolan

The aim of this project is to explore the feasibility of digital processing and computer manipulation of graphic arts quality images which are intended to be duplicated on printing presses. Specific areas of investigation include data compression, tone-scale reproduction, enhancement, input/output hardware and software, and the economical storage and retrieval of very large amounts of pictorial data.

4. DIGITAL WIREPHOTO SYSTEM

Associated Press (Grant)

Donald E. Troxel, William F. Schreiber, Samuel M. Goldwasser, John N. Ratzel

Since August 1970, we have been developing a news picture (Wirephoto) distribution system that is entirely new for the Associated Press. It is to be introduced in stages, in such a way that at least the present standard of quality and service will be maintained everywhere, with improvements spreading gradually to all locations.

The ultimate system as now envisioned will operate as follows. Pictures will be stored under computer control. An editor can then view any picture on a TV display in order to select, discard, edit, transmit, or store that image for later automatic dispatch. Editing may include cropping, enlarging, reducing, tone-scale enhancement, sharpening, combining, and addition of captions. No additional chemical photographic work will be required for any of these picture-processing operations.

Transmission over the "backbone" system linking AP bureaus and large metropolitan newspapers that have substantial computer facilities will be via high-speed digital links and will originate and terminate generally at computer-controlled digital storage devices. Transmission to subscribers will be analog or digital and at speeds and scanning standards appropriate to the existing transmission facilities. Complete control will be exercised by the New York network monitor. In the absence of manual interventions, transmission to all points among the bureaus, from point to point, and to regional networks, will be accomplished automatically.

We have implemented some of these procedures in the laboratory, using a PDP-11

computer (80k core, 38 megabit disk). The input may be a picture from the AP network, from a local analog transmitter, magnetic tape or Dectape, and is stored on a disk. Pictures may be transmitted from the disk to comparable receiving points. Pictures stored on the disk may be viewed on a TV display utilizing a full-frame semiconductor storage system. Editing facilities already in operation include cropping, enlarging or reducing, combining several pictures into one, addition of captions, and sharpening.

The multitask software operating system permits new picture-processing operations to be integrated easily, and we plan to keep incorporating additional picture-processing routines into the system.

We are particularly interested in picture-processing operations in which the processing depends on the local content of the picture. That is, the detailed parameters of a coding or enhancement scheme vary for different local areas. In this type of processing it is of prime importance to avoid artifacts such as contours outlining these local areas. We are also accelerating our interest in color picture processing, both from the viewpoint of coding for bandwidth compression and enhancement or manipulation.

The Associated Press is now in the process of installing the computer-based image processing system in New York City. When operational, it will initially be used to coordinate the newsphoto transmissions between the domestic and international Wirephoto networks.

5. SIMULATION OF PARTIAL DENATURATION MAPS (PDM) OF BACTERIAL PHAGE

Ian T. Young, Donald S. Levinstone

In studying the PDM of long DNA molecules we have made extensive use of the PDM histogram generated by algorithmic procedures.[1,2] In order to determine the significance of the histogram prototypical maps isolated in this manner, the alignment procedure was run on several sets of randomly generated data.

The histogram T6056 (Fig. XXIV-2a) is generated from 63 artificial denaturation maps each of which has the identical number and lengths of denaturation bubbles as a corresponding curve of the original P22 data, but a randomly chosen position – i.e., these data have essentially identical statistics to the real data, except that the postulated regional dependence of denaturation is lacking. The three histograms T0066, T1856, and T0876 (Fig. XXIV-2b, c, d) were generated from artificial PDM with the number of denatured regions per curve selected by a Poisson distribution having the same measured statistics as the original P22 ensemble, with the length of the region chosen
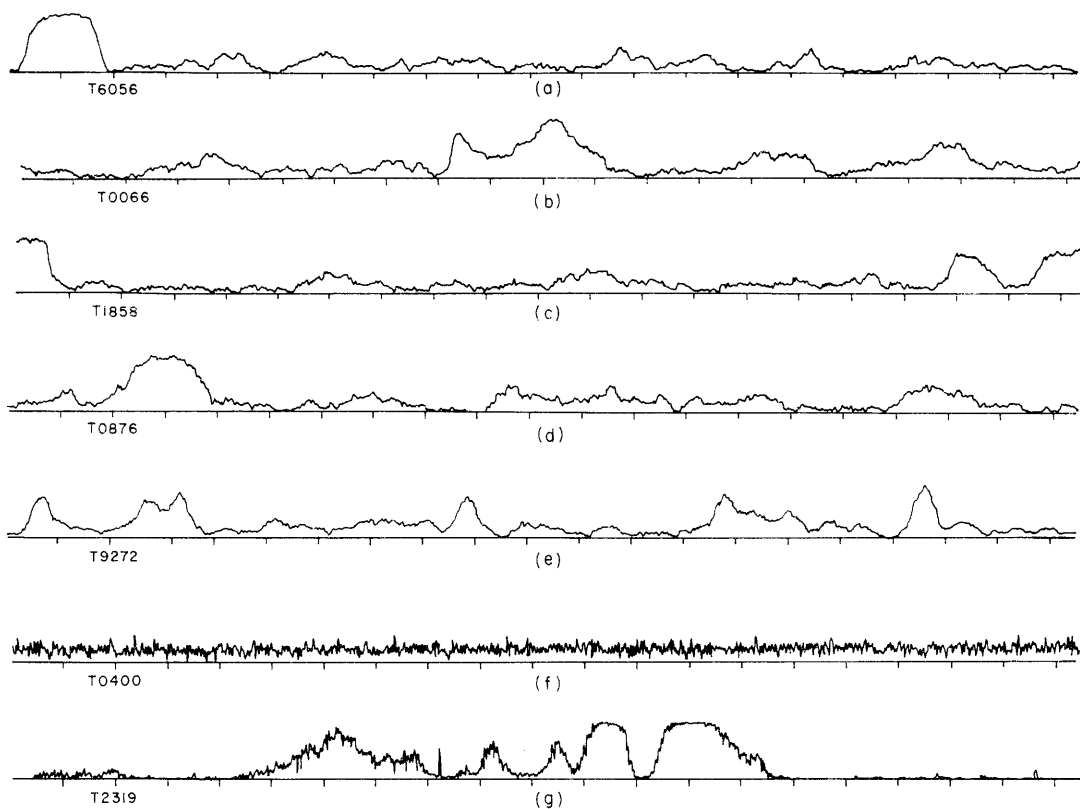
Fig. XXIV-2. Partial denaturation map histograms of stochastically
generated molecules.

via a probability mass function accumulated over the original P22 data ensemble, and
with uniformly random points of denaturation.

The histogram T9272 (Fig. XXIV-2e) was generated from artificial PDM with two
Poisson distributions, for the number of bubbles per curve and the lengths of these
regions (with parameters chosen to agree with the measured parameters of the original
P22 data), and random region positioning.

It can be seen that while the histograms converge to a peaked pattern, the average
correlation coefficient $(\bar{\rho})$, a measure of the extent to which the computer faithfully mod-
els its components, is in general quite low, as contrasted with the results on the real
P22 data, as shown in Table XXIV-1.

The conclusion that may be drawn from these simulations is that relatively consis-
tent position information (with respect to denaturation) among the curves comprising
the ensemble is an important factor in convergence to a pattern which correlates well
with its substituents. The data set corresponding to histogram T0400 (Fig. XXIV-2f)
was generated by a coin-flipping process. The probability that any individual point of
the 1024 comprising a curve will be a 1 is .21, the same as the average probability of a 1

Table XXIV-1. Average correlation coefficients for various PDM histograms.

PDM Type	$\bar{\rho}$
P22	.72
T2319	.71
T6056	.51
T0066	.46
T1858	.47
T0876	.48
T9272	.41
T0400	.15

anywhere in the original P22 ensemble. These data clearly contain no sensitivity to position, and none of the regional dependences characteristic of the real P22 data. Although these denaturation maps are essentially white noise, it is evident that a pattern can be found such that the average correlation with it is .15 rather than 0. This occurrence is due to the rotation of the curves to maximize correlation with the prototype in the alignment process. The statistics of this set of data thus establish a lower bound on the $\bar{\rho}_{proto}$ of the alignment process.

The final simulation with P22 statistics is represented by the histogram denoted T2319. In this experiment a similar coin-flipping procedure was utilized, but with the probability of 1 at any position equal to the fraction of the curves denatured at that position in the histogram of P22 data (used as a representation of the true position dependence of denaturation in P22). Thus, denaturation "bubbles" are not generated, but the overall regional nature of the denaturation process in the P22 data is mirrored in Fig. XXIV-2f. It may be noted here that the $\bar{\rho}_{proto}$ for these data is almost identical to that of the real P22 data set.

The conclusion drawn from these simulations is that the principal information recognized by our pattern-isolation scheme is the position dependence of the denaturation.

Our alignment procedure appears to be dependent on relatively consistent data, with the significant denatured regions of similar extent, for good construction of a pattern clearly different from an alignment of noise.

References

1. I. T. Young, D. Levinstone, M. Eden, B-K. Tye, and D. Botstein, "Alignment of Partial Denaturation Maps of Circularly Permuted DNA by Computer," J. Mol. Biol. 85, 528-532 (1974).

2. I. T. Young and D. Levinstone, "Recognition of Partial Denaturation Maps (PDM) of Bacterial Phage," Progress Report No. 119, Research Laboratory of Electronics, M.I.T., January 1977, pp. 116-118.

6. ANALYSIS OF CELL AGING

Ian T. Young, Donald S. Levinstone, Irvin S. Englander

This report describes an initial study toward the automatic location of fibroblast cells in a field under Hoffman contrast microscopy. Figure XXIV-3a shows an example of such a field. The goal is a set of contours outlining the cells present in the field.

The study was performed on the Cognitive Information Processing Group's PDP-9 Biological Image Processing Computer using the PIXIE interactive picture experimentor, developed by Englander.[1]

As a first step, we generated the gradient picture from the original picture by convolving the original with the following directional derivative masks:

$$\vec{H} = \begin{array}{|c|c|c|} \hline -1 & 0 & 1 \\ \hline -2 & 0 & 2 \\ \hline -1 & 0 & 1 \\ \hline \end{array} * \vec{P} \qquad\qquad \vec{V} = \begin{array}{|c|c|c|} \hline -1 & -2 & -1 \\ \hline 0 & 0 & 0 \\ \hline 1 & 2 & 1 \\ \hline \end{array} * \vec{P}$$

The gradient is the point-by-point calculation

$$g_{xy} = \sqrt{H_{xy}^2 + V_{xy}^2}.$$

The gradient is shown in Fig. XXIV-3b.

A threshold operation was performed on the gradient; the resulting binary picture was filtered using the following mask,

$$\begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline \end{array}$$

and a threshold which accepts points for which at least four of the points are in the binary picture. This new picture was "inclusive-ored" with the binary picture.

The resulting picture was contoured. Small regions were eliminated from the result, as were interior exclusions. The trace lines were smoothed in regions where a short potential perimeter enclosed a large exclusion. The final result is shown in Fig. XXIV-3c.
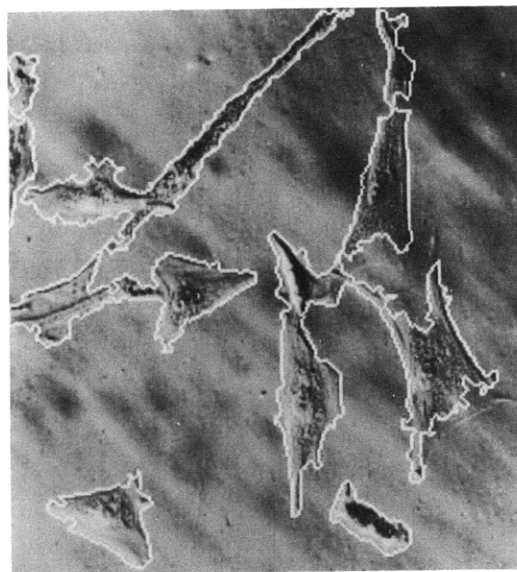
It can be seen that the results of the initial study are promising. All of the cells in the picture have been located and outlined. Two sections of the cell at center right have been

(a)
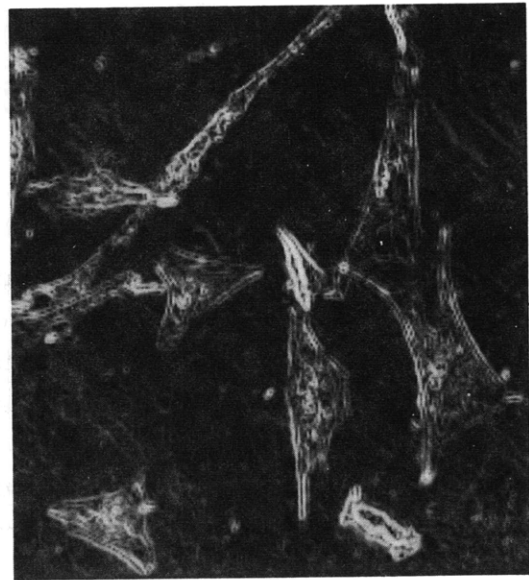
(b)

(c)

Fig. XXIV-3. Fibroblasts observed under Hoffman modulation
contrast microscopy.
(a) Original (digitized) image.
(b) Gradient image.
(c) Final contours superimposed on cells.

eliminated; although pieces of these regions existed in the original contour trace, there was not sufficient evidence under our criteria to attach them to the cell. Further research is needed in this area. Additionally, no effort was made to deal with the problem of overlapping cells, although the intermediate picture data give some indicators as to an approach to the problem. Considerable future research will be required in this area.

References

1. I. Englander, "PIXIE — An Interactive Experimentor for Pictorial Pattern Recognition," Ph.D. Thesis, to be submitted to the Department of Electrical Engineering and Computer Science, M.I.T. in February 1978.

7. TEST FOR DONOR-RECIPIENT HISTOCOMPATIBILITY

Ian T. Young, Stephanie E. Sher

The major thrust of our work on mixed lymphocyte cultures from June 1976 to January 1977 was in image analysis. A system was developed to enter a selected area of photographs of MLC wells into the PDP-9 computer. The pictures were read in through a video camera attached to the Tempo computer. The Tempo coded the pictures into levels of brightness between white and black, and transmitted the information in digital form to the PDP-9.

A student, Mark Alan Shalloway, analyzed six sets of photographs, one set from cells compatible at the LD and HLA loci, one set compatible at LHA, but not LD locus, and four incompatible at both loci. He identified two kinds of textural features which differed between wells. One consisted of large clusters of cells, some with smooth, well-defined edges and others with rough, diffuse edges. He termed this texture macrostructure and concluded from the data that it represented a response of the cultured lymphocytes to the serum in the culture medium. The other texture consisted of individual cells and the spaces between them. He termed this microstructure.

He developed an image-processing system based on analysis of the distance between adjacent dark points. This corresponded to information on cell size and distribution. Using this analysis of microstructure, the ability to predict the compatibility on the six plates was inconclusive. We attribute this to two areas: Our inability to understand how humans can successfully predict the compatibility based upon textural features[1] and the heterogeneity in the cell population from use of the human model when we are trying to determine the controlling variables.

The work will be continued under a grant from the Leukemia Foundation using a different type of image analysis and a different animal model.

References

1. See Progress Report submitted with request for grant renewal by Murray Eden, February 10, 1976; also see Progress Report No. 117, Research Laboratory of Electronics, M.I.T., January 1976, p. 100.