

2005年（第21回）日本国際賞受賞者 2005 (21st) Japan Prize Laureate



長尾 真博士（日本）

独立行政法人情報通信研究機構理事長
1936年生れ

Dr. Nagao, Makoto (Japan)

President, National Institute of Information and
Communications Technology
Born in 1936

人間的情報処理を目ざして

(1) 私が学部を卒業して大学院に入った1959年頃は、日本はコンピュータの黎明期でありました。修士2年になってようやくコンピュータが使えるようになりましたが、アセンブラ言語という機械語でプログラムを書き、紙テープでプログラムやデータを入力していた時代でした。外国の学術雑誌もほとんどなく、来ても半年以上遅れたもので、コンピュータは万能であるといった議論がなされているということが何となく分るといった状況でありました。

私は子供の頃から人間の心や頭脳の働きといったことに興味を持っておりましたので、コンピュータの万能性ということに関心を持ちました。もしそうであれば、コンピュータに最も人間的な仕事である言語翻訳をさせたり、文字の認識をさせることが出来るのではないかと考えたのでした。最初はチョムスキーの文脈自由文法から入りましたが、すぐこれだけでは言語は扱えないと分り、ロジェのシソーラスを参考にして意味素という概念を導入し、これによってできるだけ意味のある文が生成されるようにしました。これは1965年のニューヨークでの第1回計算言語学国際会議で発表

し、注目されました。

このとき重要なのは文の中心的機能はたず動詞の働きです。いろいろな文法をしらべ、フィルモアの提唱した格文法がコンピュータにとって妥当な文法であること、さらにある動詞の格フレームは、意味素によってそこに入りうる単語に制約を科さねばならないことに気づきました。

1982年から4年間、当時の科学技術庁の振興調整費によって科学技術論文の抄録部分の日英、英日翻訳システムの研究開発を行いました。そこでは格文法の考え方をとりました。出来上がったシステムは、その後辞書データが増やされたり、文法やシステムが改良されたりして現在も使われています。

(2) この研究をスタートさせながら、一方ではこの枠組のシステムの限界もある程度分ってはいました。それは例外的表現に対応させて文法規則を幾らでも書かねばならないこと、意味素の数を幾らでも増やしてゆかねば微妙な表現をうまく翻訳できないことなどでした。そこで人間はといったいどのようにして翻訳をしているのだろうと考え、次のことに気づきました。すなわち、

人は一般にある表現の翻訳を他人から教えてもらって、類似の表現をおなじように翻訳しているということでした。そこで、文や適当な長さの句表現に対応する英語表現を対訳辞書として用意しておき、類似の表現が出て来た時にその辞書を参照して“なぞり翻訳”をするという翻訳方式を1981年に考案し、これを「アナロジーに基いた翻訳」と名付けました。これは種々の新しい表現に対しても対訳辞書の内容を増やすだけで対処でき、文法や意味素などの人工的な枠組の変更といった面倒なことをしなくてすむほかに、良質な翻訳表現が得られるという利点があります。そういったことからこの翻訳方式は、今日“用例翻訳”という名称で、国内だけでなく海外でも広く使われるようになって来ております。

- (3) 文の解析が失敗する原因を調べますと、1つの文の長さが長い場合であり、そこには並列表現が複雑に使われているということがありました。そこで文中のどこどこが並列的な構造をしているのか検出できればよいわけで、その検出にはダイナミックプログラミングの手法を導入して成功しました。この成功に刺激されて、言語処理のいろんな課題の解決にダイナミックプログラミングが広く用いられるようになってきております。
- (4) 機械翻訳の研究を通じて分って来たことは、この分野の研究こそ国際的に協力して行うことが必要であるということであり、そこで、1970年代の終わりころから電子工業振興協会や通産省に働きかけ、機械翻訳の調査団を欧米に派遣したりする一方で、1986年からほぼ毎年種々の国際会議を主催し、1991年には機械翻訳国際連盟を結成し初代会長を務めました。この連盟は2年毎にアジア、欧州、アメリカの順で国際会議を開催し、機械翻訳の技術、利用者の立場からの翻訳システムの使い方の工夫等について研究発表と情報交換を行ってき

ました。国内でも言語処理学会を作り、この分野の発展を心がけております。

- (5) コンピュータに人間的なことをやらせるもう1つの課題は、画像の解析と認識であります。1960年代にまず文字認識の研究から入り、初期の郵便番号読取機に応用されたこともありましたが、多くの人が文字認識の研究をやり始めましたので、当時まだ誰もやっていなかった写真などの画像の解析と認識の研究に移りました。まず世界に先がけて人の顔写真の解析の研究を行いました。いろいろな顔立ちの人がいて、また眼鏡をかけていたりして、単純な解析では間違っただけです。そこで顔の一般的性質である目、鼻、口、あごなどの相対的位置関係を制約条件にして、たとえば口の部分の認識がうまく行かない場合には、その上にあるはずの鼻の認識をもう一度やりなおしてから、再度口の部分の認識を行うといったように、認識処理にフィードバック過程を入れたプログラムを作り、顔の認識の精度を大幅に向上させました。

次に航空写真の解析の研究を行いました。道路や畑、林、家、自動車など写真に写っている対象を認識するためには、色や形をたよりにするだけでなく、自動車は道の上にあること、日光の具合によって家や林などには影が一定方向に隣接して存在するといった種々の条件が、全体的にできるだけ矛盾なく満たされるように認識を行うことが必要であると考え、人工知能研究で行われていたブラックボードモデルを画像認識に導入して成功しました。こういった人間的な試行錯誤をするシステム、総合的に適切な結果を出すシステムは他にない特徴がありました。

- (6) 1980年代の終り頃になって文書や画像がかなり自由にコンピュータで扱え、巨大データベースが利用可能となって来ましたので、それまでの研究の総合化として電子図書館の研究を行いました。そこには図書

検索に書誌的事項を利用した検索のほかに、目次情報をうまく利用した精度の高い検索方式を考案したり、ハイパーテキスト概念の導入、機械翻訳機能やその他種々の便利な機能をもったユーザインターフェイスの開発などを行い、企業と組んで具体的なシステムを作り、1994年にはデモンストレーションするなどして、その有効性を示しました。これは世界的にももっとも早いものの1つで、内容的にユニークなものでありました。この電子図書館方式は今日あちこちで使われております。

以上のように、人間の頭脳活動にヒントを得て種々の研究を行って来ましたが雑多岐にわたる課題の解決には、このようなアプローチがますます重要になってゆくものと考え、さらなる努力をしたいと考えております。