

電算処理による禪文献研究

—その方法的模索—

沖 本 克 己

仏教の世界にもパソコンの進出は目ざましい。しかし、ほとんどの場合その主たる使用目的はワープロ、あるいはエディタによる効率的な原稿の作成という範囲に留まっているのが実際のところであろう。

しかし、コンピュータとはその名の通り計算機である。つまり数値の記憶と高速な演算をもっとも得意とする装置である。そして我々の用いる言語や文字も簡単な数値のいくつかの組み合わせで代用できると言うことがワープロ機能のAでありZなのである。

しかし、それでは我々が所有するパソコンの能力の、ほんの一部を利用してはすぎないし、我々の研究的営為に対しても貢献するところはごく僅かに過ぎないだろう。それ故、自らの研究領域や方法を分析し、よりの確にリンクさせることができれば、パソコンをもっと優秀なツールとして利用し得るのではないかと考えるのも自然なことである。

これまで、パソコン(あるいは大型計算機)を利用した仏教学研究に関しては、過去に印仏学会が三度にわたって主催したシンポジウムの成果にほぼ集約されているといつてよい。¹⁾

それら、および関連する研究成果を管見するに、その方向はデータベース(DB)構築に関するものと、その利用の方法に関するもの、いわばインプットとアウトプットの二方向に大別できる。そして、それは後述するようにさらに幾つかの範疇あるいは方法に類型化することができるだろう。とまれ、こうした試みに先進的に取り組む人々の熱意には敬服する他はない。

しかし、キカイ・アレルギーの、既に硬直した頭脳にはこれらの成果やフィールドは難解というも愚かな地平にあり、にわかには理解することも応用することもままならない。

一方、見渡せば、現在入手し得る様々な市販ソフト(多くの場合、それは企業向けが中心なのだが)あるいはPDS(Public Domain Software)主としてパソコン通信で手に入れる無料公開のソフト)などが急速な発達を示している。これらの利用の仕方によっては、もつと簡便に我々のごときものにも、そしてその仏教学の世界にも大きな可能性をもたらしてくれるかも知れないのである。

ではそれはどのような形で、であろうか。

パソコンによる成果の実例として、例えば索引類の作成という従来膨大な労力を伴った仕事がいとも手易く実現されているのは既に周知のことである。実際、大量のデータの中から特定の記号あるいは予め指定された類語を検出し、記述し、任意の方法で並べ替え、印刷するという作業はパソコンのもつとも得意とする領域のひとつである。

しかし、実を言えばオリジナルの文献データさえテキストファイルとしてコード化してあれば、あらためてそれを単純に、およそ五倍から一〇倍の膨大な量の索引として編成しプリントアウトすることは、個別の研究者やその特定

の研究にとつてはさほど必要なことではない。必要な情報はその都度オリジナルデータから容易に取り出すことが出来るからである。

かくて、我々はパソコンという秀れた道具を前にして、それがいったい何を可能にしてくれるのか、そのことを改めて問わねばならない状況に立ち至っているというべきであらう。それを言い替えれば、パソコンやそのデータを有効に活用させるために、自らの学問は如何にして成り立っているのか、その方法論に対する反省を要求される羽目にもなっているのである。

即ち、我々の行なってきた研究は、それが単なる恣意や確執に依存するものでないならば、どのような説得力のあるプログラムで処理されているのか、ということに対する洞察も必要となってきているのである。

そして、そのプログラムの出来栄がそのまま得られた成果の持つ説得力にもなるのであるから、敢えて極論するならば、我々の学問はどこまで数式化できるのか、ということに他ならない。そして、それがつまり客観的な合理妥当性を持ち得るかどうかということと直接関わっているのである。

いわばパソコンという有能にしてかつ愚直な道具は、我々の学問の位相を改めて問う役割を同時に担っているのである。

こうした点に関して、十分な見通しもなければ、自らの学問に対する方法的反省もなお適切を欠いたままであるが、しかし、パソコンによって何が可能であり、何が不可能(に見えている)かを、従来の乏しい経験と見聞によってここで確かめておきたいというのがとりあえずの本稿執筆の動機である。

仏教学とは文献学である。如何に独創的な思想的営為であれ、文献的裏付けを欠く議論は充分な説得力をもち得ないのは自明である。そして我々が取り扱う文献にはおよそ三種類がある。ひとつは直接の研究対象としての文献資料(テキスト)であり、ひとつは辞書などのいわゆる工具類であり、今一つは関連する研究領域に関する先行論文である。我々はこのテキストとツールと研究史という三種類のDBを手に入れることによって、有効な研究を行ない得るのである。

もっとも仏教学の方法も多岐にわたる。例えばここに現代仏教学の研究方法として、同時代的な様々な調査やその統計的処理を持ち込むなら、それらはまた別のDBのジャンルを形成することになるであろうし、さらに異なった斬新な研究方法も可能かも知れないのである。しかしそのことは、ここではひとまず除外しておくことにしたい。

さて、上述の三種類のDBはいずれも周到で豊富であることが研究成果の質を左右する。それ故、これらに関する蓄積されたデータの量が重要であるし、適切な検索作業が必要である。そして、それがコンピュータのもっとも得意なジャンルであることも容易に想像できるだろう。

実際に、これらのデータ蓄積は例えば日本印度学仏教学会を中心として構想されていて、特に論文DBの方は国内の関連資料はほぼキーワード入力を終えている。やがてこれらを集約して収納する大型計算機センターに電話でアクセスすることによって、必要な情報を入力することが可能となるだろう。

しかし、実はキーワードだけではなく、論文そのものがそのままデータとして保存されるのがあるべき姿である。

それを組み合せ検索によって捜査し、そこから必要と思われる論文を容易に引きだし得る状態になるのが理想なのである。もともとキーワードの設定とコード化には途方もない労力と時間と資金が必要だったうえに、キーワードが適切でなければこの方法によるのみでは充分な検索も不可能なのである。しかもそのうえ、仮に論文の所在が分かっていたとしても、今度はそれを容易に入手できない、という悲劇的な状況には再々出会っているのであるから。

しかし最近の新聞記事によれば、この問題もほぼクリアされる見通しがついたようである。即ち、一九九一年八月六日付け朝日新聞によると、大規模なDBを構築し高速検索を容易にするためには、従来は各文書にキーワードや分類項目を設定する作業が不可欠だったが、これを不要にする検索システムの回路のLSI（大規模集積回路）化に、松下電器・情報通信東京研究所が成功した、というのである。

かくして、従来困難とされていた大量の論文そのものの直接のDB化の目処がついた。このことはさらに論文のデジタル化、つまり原稿のパソコン通信あるいはフロッピーによる提出や、同じ方法での校正作業、そして電算写植による出版という出版業務の新たな方向をも加速するであろう。これらの流れは編集作業の簡便化、迅速化ならびに大幅な出版経費の削減を可能にし、そればかりか、論文DBも同時に出来上がるというメリットまで付随するのである。

近年、仏教学の隆盛は目ざましく毎年発行される論文の数は膨大なものにのぼる。少し油断をすればたちまち研究の最先端を見失ってしまうし、異なった分野に関しては論文執筆の都度、新たな関連領域を改めて参照しなければならず、適切な研究史の概観に関する悩みは一層深刻である。

こうした時、最新の論文DBの検索が容易になるのはまことに有難いことである。そしてこのことは、逆に研究史の見落としが許されないという、いわば当然の状況も一層厳密に要求されることになるだろう。

しかし、豊富なデータは、また却って不便を引き起こすことにもなりかねない。例えば『印度学仏教学研究索引』

(山喜房仏書林、一九八七)は、コンピュータを利用して作成された、『印度学仏教学研究』一卷から三二巻迄の索引であるが、その内の例えば「浄土教」や「天台宗」や「禅宗」、あるいは「親鸞」や「道元」などの項を試しに引いてみれば良い。何頁にも渡る膨大な数字の羅列は、それを参照する意欲を失わせてしまうであろう。

幸い、これには印仏学会より頒布されているフロッピーがあり、それを用いてそれぞれに組合せ検索をかけることによってより適切な論文を発見することが可能となるのだが、こうしたことから、DBの利用には蓄積の他に排除という作業が重要なこと、そしてそれらを可能にする研究の見通しと方法が必要である、ということが明かとなる。そして、パソコンは不要のものを排除するという点でも卓越した能力をもっているのである。⁽²⁾

このように、豊富なデータと的確な検索手段、いわば蓄積と排除の組み合わせによって研究の準備段階が整えられるのである。しかし、そうした点からも年々更新される業績のトレースとそれのDB化の継続、そして過去の業績もOCR (Optical Character Reader 光学的読み取り装置) などによってDB化しておくことが必要とされるに至っている⁽³⁾のである。

そしてこのことは、当面する酸性紙問題や資源問題、さらに身近なところでは我々の貧弱な書庫の限界という即物的課題を幾分かクリアするためにも必要なことだといえよう。

なお、先に見たDBのうち、研究工具として各種の書誌目録などの資料や辞書などについて簡単に触れておこう。これらの必需品もまた我々の貧しい書齋や書庫を圧迫すること大なるものがある。それ故これらの書籍類、即ち『仏書解説大辞典』やその近代に至る成果の増補版、各種仏教辞典や各国語の辞書、そして索引などがCD-ROM (Compact disk に搭載された Read only memory 読み出し専用記憶装置) つまり電子出版化され、さらにはそれがジャンル毎に統合された形で提供されるのが理想であるが、これは我々の仕事というよりも出版事業レベルのことがらに属する

のでここでは希望を述べておくにとどめよう。

三

さて、以上のように論文DBは、これまでともかく一定の成果を挙げている。勿論、問題は残っていて、例えば『印仏研索引』は三二巻二号（昭和五九年三月）までしかなく既に時代に遅れているし、他の成果も未だ公開されていない。また海外の研究文献に関しても有効なオンラインDB化などの方策が企画されていないのが現状である。

しかし、それにも増してテキストDBの構築の方面はその準備が遅れているようである。個別に研究機関あるいは個人が必要に応じて作成しているのを耳にするが、その情報は乏しい。著作権問題や作成のためにかげられた労力の評価など、それらの持つ様々な課題を処理し、併せてそれらを合理的に統括し管理する方策が未だ構じられていないのも原因のひとつだろう。

確かにDB構築には膨大な経費と労力を必要とし、しかもいったんリリースすると著作権なんぞはなきに等しく、たちまち海賊コピーが蔓延するというのが実際の状態である。

筆者もかつて『俱舎論』などのテキストDBを作成し、関連研究機関等に無償で配布したことがある。しかし、これも相互の情報交換の糸口にさえならなかったのみならず、回り回って、とある友人から『俱舎論』のDBを内緒で分けてやろうか、と耳打ちされる始末であった。もってその閉塞した状況を知るべきであろう。そのことは余談として、願わくはこうした作業情報を集約し公開する情報センターの設置されんことを。それが有効に機能することによって、テキストDBの形成と集積は加速されるであろう。

話は前後するが、ここで入力の問題を概観しておこう。この問題を抜きにしてはDBの蓄積もあり得ないからである。

仏教の世界は言葉の洪水である。ざっと見回したところで文献資料にはサンスクリット語、パーリ語、チベット語、中国語、そして朝鮮語、日本語があり、それぞれの時代と場所による変容があり、中央アジアを始めとする周辺地域にはさらに多くの関連する言語が存在する。また、論文資料には世界各国の言語が混在し、その全てを渉猟するのはほとんど不可能といってもよいだろう。

そして、テキストDBの対象に限っても、それぞれに独自の字形と文法を持っている。そのうえサンスクリットやパーリは長母音などの特殊な文字やサンディ(文字の結合)と語尾変化があり、入力も検索も簡単ではない。

そのことについて少し考えておこう。しかもこれは決して我々の目的とする漢字の問題とも無縁ではないのである。⁽⁵⁾

パソコンにはワープロというもっぱら入力面を受け持つ機能と、エディタあるいはカードという作業面を受け持つ機能がある。そして、いずれの場合にしても、ナーガリー文字等の特殊文字に関してもそれをそのままフォント(文字のセット)として画面あるいはプリンタ上で、特殊文字を実現し利用しようとする方向と、ローマナイズしたものを利用すれば充分だという方向とに分かれる。そのことについて最初に簡単に触れておくと、パソコン導入の初期にあたる一九八〇年代前半はほとんどがローマ字入力で、キーボード上のアルファベットに工夫を加えることで(例えばaの長母音はAで代用する、などといった方法)問題も回避もしくは解決していた。⁽⁵⁾

この方法は有力な手段として現在も様々なルールで行なわれている。そして、その目的はサンスクリット、パーリのキーボードによる入力とインデックス作成、という方面に専ら向けられていた、といってもよいだろう。

こうした状況に対して、東京外国語大学の町田和彦氏は、従来バラバラであったテキストファイルの規格の統一を

提唱し、またナーガリーなどへの翻字システムを開発された。⁽⁶⁾ また関連する細かな、しかし重要な問題が着実に解決されている。例えば、梵語は一見整然と配列されているかに見え、事実その通りなのだが、例えばアヌスヴァーラなどは発生論的な問題があつて、辞書などでは複雑な位置関係を示す。(日本語の「ン」などもかつては同じ位置付けにあった。) それらのソーティング(分類、配列)処理の方法(アルゴリズム)について考察を加え、⁽⁷⁾ また、検索上やつかいなサンディや語尾変化の処理方法の解決がはかられている。⁽⁸⁾ こうした細かい部分の着実な解決は今後も要求されるし、そうした成果の積み重ねが従来の文献学的研究に一層緻密な視点を提供するものでもあるといえよう。

なお、これらの成果の上に立って、その翻字システムを実現したソフトは現在〈CAPTURE〉(チャトゥル)と名付けて公開されている。その優れた操作性は評価も高く、⁽⁹⁾ またこのソフトを用いて作られたのが、本学講師・平岡昇修氏による『サンスクリット・トレーニングⅠ、Ⅱ』(世界聖典刊行協会、一九九〇、一九九二)である。

これらはいわば手書き入力方式で、キバンチャーは梵語に関わる知識を持つことを要求される。そうしたこともあつて、このソフトはデータ構築よりも、作者も言うように論文やテキスト作りに適しているといつてよいだろう。

なお、サンスクリット、パーリ語のローマナイズ表記における特殊文字を半角文字で直接入力できるドライバ「汎」⁽¹⁰⁾も開発され、現在、印仏学会より頒布されている。その特徴はナーガリー文字はさしあたり不要という立場から、入力の便宜、利用の簡便化を目的としたもので、将来の機能拡大にも耐える能力を持っている。そして勿論、フォントさえ揃えればナーガリー文字などへの変換も容易である。

また、東北大学ではパソコンと汎用大型コンピュータをリンクしてデーヴァナーガリーの自動解読処理を行なっている。対象となっているテキストは今のところ各種テキストが豊富でそれだけに課題も多い『法華経』である。⁽¹¹⁾そして、これらの手法はテキストの自動解読にも発展し、既に成果をあげつつある。⁽¹²⁾

これらはさらに手書きを含む様々な書体の解析にも展開し、また、使用文字例や出現頻度、文章の構造的特徴など

のいわゆる計量文献学的方法への模索が進められている。⁽¹³⁾

またチベット文字に関しては、古坂絨一氏が梵藏漢の三種の文字(但し梵語は半角ローマナイズ)を混在させたテキストの形成と処理のデモンストレーションを行なっている。⁽¹⁴⁾ これらはマッキントッシュ(MAC)を使用したもので充分実用に耐え得るものである。

チベット文字に関しては、先の古坂氏の紹介するようにMAC用フォントが発売されている他、最近、PC98用のソフトが言語情報研究所により開発され、「チベット語文献処理総合システム」として発売されている。⁽¹⁵⁾ これは「同時多国語処理システム」の一部を構成し、対応言語は平成三年八月現在二〇ヶ国語を超え、処理能力にも秀れたシステムである。

また、同研究所はACID(The Asian Classics Input Project)と協同してその日本事務局となり、現在、チベット語仏教文献約五〇種類がDBとしてリリースされている。⁽¹⁶⁾

なお、チベット語のトランスクリプション方式は、全く統一が取れていなかったが、特殊記号を排したスクリプト(北村方式)が定着しつつあり、キーボードによるローマ字入力は極めて容易である。⁽¹⁷⁾

四

次に漢文文献の問題に移ろう。漢文仏典あるいはその周辺資料は他の言語に比してその分量は圧倒的に多く、また研究者の数も多数にのぼるからDBの量もかなりのものと予想される。しかし、先に述べたような事情もあり、それに関する情報量は極めて少ない。それ故、ここでは知り得た範囲に限り私見を述べるとどめよう。

先般、仏教大学で開催された印仏学会で清水光幸氏が「コンピュータ大藏経」に関する有意義な発表を行なった。

(このプログラムには本学のウルス・アップ氏も技術協力を行なっている。) 具体的にはテキストをイメージスキャナで入力し、それを「採字帳」と名付ける市販もしくはPDSソフトで、MS-DOSのテキストファイルに変換する方法である。従来のOCRによる方法に較べてその変換効率も格段に高くなっている。⁽¹⁸⁾ (なお、清水氏は同時に一字索引を作成するための優れた方法も紹介された。) これによって従来個別にかつ閉鎖的な環境で行なわれていた文献DB構築はかなりの進展が見られるであろう。

しかし、問題がないわけではない。ひとつは版權の問題であり、ひとつは欠字の問題であり、ひとつは校訂と精度の問題である。それらについて卑見を記しておくことにしたい。

まず出版物に関する権利には著作権と出版権があるが、我々は従来こうした方面には暗く、十分な知識も持ち合わせてはいない。しかし、研究という大義名分がすべてに優先するわけのものでもないのであれば、最低限の理解は得ておくことが必要であろう。もっとも、文献DBの著作権ということになれば、それは仏陀もしくはいずれも名を残していない經典制作者、あるいは翻訳者に帰属することになる。論典もまた同じ事である。そして保護期間は既に過ぎているし、經典は自ら書写読誦を勧めるものであるから、ここには著作権は存在しないと考えてよいだろう。もっともこれは半分冗談として、印刷された資料を原典としてDBを構築する時、著作権とは別の出版権(版面権)の問題に関わることは明確にしておく必要があるであろう。

例えば、大正大藏経は国際的にも声価が高く、その利便性からも今後ともその資料価値を減じないと考えられるから、DBの基本テキストとして既に各方面で利用されてきている。そして、それを原典として手間のかかるキーボード入力をしている時点では、その膨大な作業量からしても、また「研究目的」という理由付けからも、あまり版權との関わりを意識する必要はなかったように思う。しかし、OCRなどによる画像データとしての取り込みや、それを

もとにしたデータ変換などの方法は限りなくコピーに近く、版面権を侵しているおそれは充分にあるというべきであろう。

しかも一方で、コンピュータ入力も完璧なものではなく精度はさほど高くはない。そして仮に技術の進歩や綿密な校正によって限りなく完璧なものが出来たにしても、使用中に誤入力や破壊が起る危険には常に晒されているから、それに依存しきることはできないのである。そして、テキストDBが常に抱えるこうしたある種の不完全さが、逆に版權問題をもクリアする根拠となり得るかもしれないのである。

また、このテキストDBの構造的な不完全さを示す事例として欠字の問題がある。即ち現在パソコンで使用し得るJIS漢字は第一水準が二九六五字、第二水準が三三八八字にすぎない。一方、もっとも秀れた漢字辞書のひとつである諸橋轍次『大漢和辞典』(大修館)に収録されている漢字の総数は四八九〇二字である。しかしこれとて完全ではなく、異体字や仏教特有の音写文字などを勘案すればその数はほとんど無数といつてよいだろう。つまり、漢字数には妥当な限界というものがなく、しかも利用可能な字数をとにかく増やすことが必ずしも効果的な方法でもないのである。かくして漢文にあっては欠字問題は避けることが出来ぬ欠点であり限界となるのである。

これを回避するために例えば『大漢和辞典』の漢字番号を用いる方法(清水光幸氏)や、半角アルファベット二文字の組み合わせで処理する方法⁽¹⁹⁾、など様々な方式がそれぞれに工夫されている。しかしいずれも便宜的な回避策でしかないことには変わりはない。

このように、DBの構造的な不安定性や欠字、そして単純にディスプレイの視界の狭さやウィンドウの限界など、様々な要因が重なって、大正大蔵経を原本とするDBはあくまで研究の補助的な道具として位置付けられることにならざるを得ない。原典にとって代ることは恐らくあり得ないのである。たとえ大正大蔵経が全てCD-ROM化されたとしても。だから大正大蔵経そのものが不要になる事態も考えられず、むしろその必要度は高まると考えてよいだろう。

う。それがさしあたっての私の見通しである。大型コンピュータを導入したオフィスが当初の予想に反して、ペーパーレスになるどころか、却って紙の洪水に巻き込まれているのもあるいはその一例かもしれない。

最後に校訂の問題に触れておこう。これは作成されたテキストDBの精度とも密接に関わる問題である。

大正大蔵経が、その普及率からいっても今後とも原テキストのスタンダードを形成していくのは確かである。しかし、例えば禅宗学の研究で、便宜性の問題を離れてそれをそのまま一次資料として使用するものはまずいだろう。原典としての大正大蔵経そのものの正確さ、つまり誤植と校訂の問題があるからである。そしてこの校訂という人為も常に錯誤を呼び込む結果になる上に、ある面では極めて恣意的で主観的な読み方を持ち込む危険性に常にさらされているのである。

かくして我々にとって最良の方法はオリジナルテキストそのものの直接の参照という、ごく当たり前の結論にたどりつくのである。そして、その時パソコンに出来ることといえば、それらを画像データとして提供することであるが、これは膨大な記憶容量を必要とする上に自由な処理は出来ないから第二義的な用法に留まるだろう。

以上のように漢文DBには様々な制約や限界があるが、それがDB構築の妨げになるというものではない。むしろ、こうした位置付けを明確にした上で、DBの総量を増やすことは有効であり必要なことである。そして、その上で本論の狙いでもある、それによって何が可能であるか、が検証されるべきであろう。

五

以上をまとめれば、各種のDBは、ROMの普及による重量級の辞書などの書齋からの排除と、大正大蔵経などの

原典資料の利用時における補助的ツールとしてのテキストDBの位置、と用途に依ってその性格や書籍との関係が異なることに気がつくであろう。

つまり、辞書などのツールや書誌DBはひとまず別にして、テキストDBが従来の書籍にとってかわるとは予想し難いのである。しかしまた、本がこれからも本であり続ける必要もない。必要に応じて離合集散する、そうした、本に関する新しい視点もまた増えて行くことだろう。

最後にアウトプット、つまり研究成果として如何なることがパソコンによって可能であるかについて考えてみたい。パソコンによる研究成果として、まず誰もが思い付くのが索引や辞書の作成であろう。その性能および性質からして検索の方面に卓越した機能を発揮するのは先に見たとおりである。それ故、語彙や用例の研究にはDBはもはや欠かせぬし、それがこれから先、質量ともに長足の進歩を遂げるのも確かだろう。

むしろ、いままで苦勞して渉猟したデータが、今後は集まりすぎて整理に追われるという破目になるのも既に現実である。精密で網羅的であることは必ずしも有効なことではないのである。そしてまた、一字索引のようにシステム化され単純作業化された仕事は、翻訳と同じ様な意味で、これはもはや研究成果とは呼び難いだろう。

かくて、例えば索引に関して言えば用語索引、それも諸種の関連するテキストを合糅した語彙索引といったものが有効な方法ということになるであろうか。たしかに関連資料を網羅した術語索引はその需要を増すと考えられる。つまり機械的な一字索引よりも確な語彙索引の要望が強まると考えてよいだろう。それはつまり、整理された情報の書籍による提供ということである。そしてそれは、いちいちパソコンに頼るよりも効率的である、ということを示しているのである。

禅籍のDB化に関して既に秀れた実績のある禅文化研究所では、そうした発想に沿って、『禅林句集』や無着道忠

の『虚堂録翠耕』などを、影印判テキストと索引の組み合わせで出版している。これらはさらに基本典籍叢刊として続刊されるようだが、願わくば索引がフロッピーで提供されんことを。そして、更にそれらを合糅した語彙索引が廉価でリリースされんことを。そうすれば我々はそれを任意に組み合わせて一気に検索の実を挙げることが出来るのである。以上のように索引そのものにも様々な方法があり得るのだが、共通して言えることはそれらはツールの一分野を形成するものである、ということである。

では次に、論文の中核をなす創造的な部分についてはどうであろうか。残念ながら、企業などで用いられているソフトを見ても、総合判断や意志決定といったトップレベルのブレインワークにはコンピュータは関与し得ないのが実際のところである。

そのことは論文においても同じことで、独創的な部分は少なくとも現段階ではパソコンには処理できないし、将来にわたっても、そう簡単ではないということは言えそうである。

それはアルゴリズム(計算の手順)についても、パソコンは呆れるほど単純な計算の反復によって作動している。それに対して我々の思考は非常に複雑でかつ「跳躍」が多く、俄かにはそれを算式として表現できないのである。

また、人文科学では同一のデータ、同一の資料に対して、必ずしも同じ結論が出てくるわけではない。このあたりに、人文科学の限界と広がりがあるといってもよいかもしれないが、しかしこのことは自然科学にあっても事情は大して異ならない。例えばユークリッド幾何学と非ユークリッド幾何学のごとくである。地球上では、とまでいわずとも、球面上では「平行線は必ず交わる」のである。またニュートン力学とアインシュタインの相対性理論との関係のように。そして身近なところでは我々は相変わらず天動説をとっているように。

つまり、スケールの置き方によって我々に見える事象は異なるのである。したがって、データとプログラムがあれ

ば常に普遍妥当性のある結論がはじきだされる、と思うのは切ない初夢に過ぎないし、仮にそんな時代が到来したとしても、その処理は人間の手によってのみなされるのである。

禅宗学の研究方法は歴史的研究と思想的研究、そして文化的方面の研究とに分類し得るであろう。⁽²⁰⁾ これらはさらに時間的、地域的範疇によって、またそれに対峙する主体の姿勢によって様々に分化し、また結合するのであるが、その実例として、例えば『禅学研究』六一号（一九八二）以来、西尾賢隆氏によって「禅学関係雑誌論文目録」が連載され、秀れたDBが構築されている。

それを通覧すれば、全体の九割以上が人物あるいは書籍、もしくは語彙、概念の研究で占められている。そのことの当否は別にして、こうした文献学全盛の状況にあれば、パソコンによるデータ処理はまことに目ざましい働きをすることが出来るだろう。

また、思想表現の数量的計測、つまり計量言語学の方法による計量分析の応用も著作の真偽問題や成立時期推定に有効な働きをするであろう。

そして、こうした試みも既に行なわれているようであるが、管見するところ万人の納得する結論にはいまだ遠く、ひとつの有力な意見に留まっている様である。何よりも同一人物であったにしても、文体には揺れがあるということ、つまり、基礎資料そのものの「不正確さ」という単純なルールでは測り切れない問題に我々の学問は絶えず直面しているということがその困難さを増幅しているのも確かなことであろう。

以上、はなはだ雑駁な駄弁に墮した感があるが、パソコンに関する現状と問題点を気がつくままに羅列してみた。もともとパソコンそのものに深い関心があるわけではないので、情報に遺漏も多いかと思うし、技術的な方面に關す

る無知は覆うべくもないだろう。諸賢の御教示をまちたい。

最後に、本学国際禅学研究所ウルス・アップ氏の個人的努力により発行されている「NEWSLETTER」はこの方面に關しても有用な情報に盈ちている。そのことを記してともかく本稿を閉じることにする。

- (1) 第三七回学術大会(東京大学、一九八六)、第三九回大会(北海道大学、一九八八)、第四一回大会(東北大学、一九九〇)におけるシンポジウム。それらの総括はそれぞれ、江島恵教「シンポジウムへインド学仏教学におけるコンピュータ利用」について「印仏研三五―二、一九八七、八五六頁以下、今西順吉「シンポジウムへインド学仏教学におけるコンピュータ利用の現状と問題点」について「印仏研三七―二、一九八九、七九五頁以下、磯田照文「特別部会へインド学研究における科学技術の応用」について「印仏研三九―二、一九九一、八四三頁以下においてなされ、同時に発表者による報告およびその概要などが掲載されている。
- (2) 排除に加えてもうひとつ、いわば検索作業のバリエーションに過ぎないのだが、不存在の証明、つまり該当する範囲の中に任意の語彙などが存在せぬことを明確にする作業も造作無くこなしてくれる。大室照道『「ミリンダパンハ」の研究』「印仏研三〇―一、一九八一、一二〇頁参照。
- (3) 継続的な作業の追加刊行は従来「補遺」という形で行なわれてきた。しかし、これが利用にはすこぶる煩雑であることは周知のことである。この点でもパソコンのD Bは秀れた特性を持つといえよう。
- (4) D B構築については様々な提言がなされている。比較的初期に属するものとしては以下のものがあげられる。大室照道「仏典研究におけるコンピュータの効率的利用」印仏研三三―一、一九八四、一二〇頁以下、筑後誠隆「仏教学におけるコンピュータ利用の諸問題」印仏研三三―一、一二二頁以下。
- (5) 田端哲哉「インド仏教学とコンピュータ」印仏研三二―一、一九八三、四三三頁以下。
- (6) 町田和彦「インド系文字のコンピュータ用翻字システムについて」印仏研三六―一、一九八七、三二九頁以下。
- (7) 町田和彦「パソコンを利用した語彙研究」印仏研三八―二、一九九〇、八〇〇頁以下。
- (8) 同「文字列置換について」同三九―二、八三三頁以下。
- (9) 金沢篤「パソコン時代のインド学」駒沢大学仏教学部論集二二号、一九九〇、四〇四頁以下。その使用経験が記されている。
- (10) 金田康正、他「インド学仏教学論文データ・ベースの作成・検索・表示システムについて」印仏研三七―二、

- 一九八九、八〇五頁以下。
- (11) 塚本啓祥「コンピュータによる仏教混淆梵語の研究(1)」印仏研三七―二、一九八九、八九九頁以下。川添良幸「同(2)」同八九一頁以下、山崎守一「同(3)」同八八四頁以下。
- (12) 小島正美、他「チベット文献自動認識について」印仏研三九―二、一九九一、八四八頁以下。
- (13) 川添良幸、他「梵文法華経に現れる各種字体の比較研究」印仏研三九―二、一九九一、八三〇頁以下。同「科学技術の利用による仏教研究」同、八二八頁以下。
- (14) 古坂紘一「パーソナル・コンピュータによる梵蔵漢仏典の対照」印仏研三九―二、一九九一、八三八頁以下。
- (15) 言語情報研究所「一四二東京都品川区小山四―九―一五」
- (16) The ASIAN CLASSICS INPUT PROJECT Release 1 Release 2, ACIP 日本事務局, 1991.
- (17) 北村甫、長野泰彦『現代チベット語分類辞典』汲古書店、一九九〇、xviii. 同書はその資料分析や情報処理に国立民俗博物館の電算機を活用している。なお、梵蔵漢学においても注目すべき試みが行なわれている。神子上恵生他「コンピューター使用による仏教文献の研究(1)」、(2) 仏教文化研究所紀要二四集、二五集、一九八五、八六、参照。
- (18) 従来のOCRによる方法は設備費が高い上に、精度は八〇から九〇パーセントにとどまり、実用的とは言いがたかった。清水氏のは優れた辞書を用いれば精度は九九%をこえる。もつとも、九九%とは大正大蔵経で言えば、一頁におよそ二〇字の誤植があるということを意味し、そのままでは実用レベルと言えないのは当然である。
- (19) 西村恵学「黒豆とパソコン」基本典籍叢刊月報#〇〇〇、禅文化研究所、一九九一。
- (20) 拙稿「禅宗(臨済宗)」平川彰編『仏教研究入門』大蔵出版、一九八四、二八九頁以下。