

講演

人間の情報処理過程の数量的解析*

印 東 太 郎**

1. はじめに

つい大それたテーマをかかげてしまったが、考えてみると、数量的解析が比較的進んでいるのは聴覚と視覚における感覚的情報処理過程についてであって、より高次の言語的情報処理過程の研究は未だ余りサマをなしているとはいえない。しかし、読者が興味をもたれるのは、人間の記憶、推理、問題解析などの解析であろう。語尾変化などを除いて数えると、教育を受けた英語国民の語彙は約 75 kW と推定されている。日本語でも余り違わないオーダーであろう。会話を行い、文章を綴るには、瞬時にその中から必要な単語を取り出す IR が行われているに違いない。この場合、日本語からそれに該当する外国語をレトリブするような場合を別にすれば、キイ・ワードに相当するものではなく、何がキイになって検索が行われているのか自分でもよくわからないであろう。

筆者は、最近、人名など固有名詞を突如としてレトリブできなくなるという不幸な兆候を経験するようになった。講義のように構えてしゃべっている時よりも、雑談している時の方がおこり易いようで、先日もアセンブラという言葉が出なくなり、全くアゼンとした。そして、口惜しいことに、少し時間をおいたら、ボンと自然に思い出された。こういう tip of tongue, ノドまで出かかっている状態においては、何かを指向していることは自分でも感じられ、ある響きの言葉とか、第1音がアであるとか、漠然とした印象の存在することが多いが、検索が正常に進行している間はこういうキイも意識されないであろう。以下、不十分ではあるが、人間が自然言語を用いる場合、その根底にあるであろう心理過程に少しでもつながるような実験について若干述べることにする。

2. LTM と STM

大別すると、人間の記憶は長期記憶 (long term memory, 以下 LTM) と短期記憶 (short term memory, 以下 STM) とに区別できる。前者はほぼ一生の間ついてまわる記憶を指す。上述の 75 kW の語彙は LTM におさめられており、その他に、自分の誕生日はこれこれ、本籍地はしかじかなどぼう大な個々の事象の記憶もおさめているであろう。一方、必要な間だけ一時的に記憶し、不用になれば直ちに忘れられてしまう情報も多い。これを STM と呼ぶことにする。研究者によっては、極めて短時間しか耳に残らないような記憶だけを STM と称する人もあるが、ここでは STM をもっと広い意味に用いることにしたい。

人間の記憶、忘却に関する研究は実験心理学の誕生と同時に始まり、有名な Ebbinghaus の忘却曲線は 1885 年に発表されている。無意味綴りのリストを暗記し ($t=0$), t 時間経過してから未だ覚えている分量 $R(t)$ を測定すると、 $R(t)$ は $t=8$ 時間位まで急激に減少し、 $R(0)=1$ というスケールで表わすと、 $R(8)=0.36$ 、以後、忘却の進行は極めてゆるやかで、 $t=30$ 日になっても $R(30 \times 24)=0.20$ 位は残っている。暗記する材料として、相互に関連のない単語のリストを用いても結果は変わらない。筆者と久野¹⁾とは、こういう言語材料ではなく、ピアノの弾奏の記憶について実験を行ってみた。与えられたパッセージを $\tau(0)$ 分かけて練習し、つかえずに弾けるようになってから、 t 分後に再び弾かせ、そこでまた誤りなく弾けるようになるまでに要する練習時間 $\tau(t)$ を測定し、 $R(t)=\{\tau(0)-\tau(t)\}/\tau(0)$ とおくと、 $R(t)$ は Ebbinghaus と全く同じ型になる。ただし、こういう筋肉的運動の記憶ははるかに速やかに忘却され、 t は分で表わされる (図-1 (次頁参照))。いろいろな微分方程式を解いて $R(t)$ を導くこともでき、図-1の実線と点線は $R(t)$ に関する二つの理論曲線の比較を示したものであるが、ここで

* 情報処理学会第 16 回大会、招待講演 (昭和 50 年 11 月 22 日)

** 慶応義塾大学教授 (心理学)

慶応義塾大学情報科学研究所所長

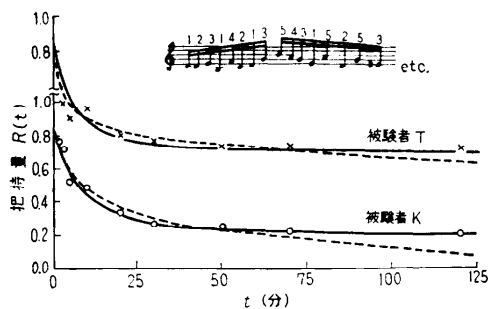


図-1 忘却曲線の例

はこれ以上立ち入らない。

上例のような実験はすべて STM に関するものといっ
てよい。ウサギ、アマダレ、キセル、……というよ
うな相互に無関係な単語からなるリストを暗記したと
して、 t 時間後にアマダレを思い出せなくなったとし
ても、それは LTM にあるアマダレという言葉のを忘
れた訳ではない。ただ、このリストにその言葉があっ
たことが忘却されただけである。無意味綴り、ピアノ
の弾奏についても同じことがいえる。実際、実験は STM
の方がはるかにやり易いので、従来の心理学的実験
の大部分は STM について行われたといっても過言で
はない。例えば、前述の“ど忘れ”、tip of tongue は
LTM からの検索に関し何かを示唆してくれるはずで
あるが、これを実験的に再現することは難しい。

正月も近いが、百人一首というゲームを考えてみよ
う。上の句と下の句を記憶しておくのは LTM の機能
である。しかし、選手級になると、その実力は STM
の有効な利用にかかってくるらしい。

まず、第一に自分及び相手の札の配列を記憶しなけ
ればならないが、これは STM の機能であろう。全日
本かるた協会競技規定によると、50 枚が空札になり、
25 枚ずつ相手と分けるのだそうである。選手達は、
自分の 25 枚に関する限り、並べ終わった時にはその位
置は STM におさめているということで、それは自分
の配列のシステムが確立してはじめて可能になる。
このシステムの記憶は LTM であろう。

記憶の実験において、ウサギ、アマダレ、キセル……
など与えられたリストを覚えられるのも LTM に存在
するシステムのサポートであるからで、日本語を知ら
ぬ者に簡単に記憶できる訳がない。われわれにとって
外国語はゆっくりしゃべって貰わなければ聞きとれな
いの、かるたの場合、相手の札の配置の記憶に骨の
折れるのも、LTM にそのシステムがないからで、

STM の容量は LTM のサポートによって変化する。
ともかく、かるたの場合、彼我の札の配置の STM にお
ける記憶は常に up date されていなければならない
であろう。なぜ記憶する必要があるかは後にふれる。
第二に、競技の経過も STM におさめられていなければ
ならない。各札はいわゆる“きまり字”をもち、小
倉百人一首全体としては、“君がため”は次が“春の野
にいで”の「ハ」とくるか、“惜しからざりし命さえ”
の「オ」とくるかによって枝かわれる。しかし、“君
がため”の一枚が既に出てしまうか、配置された 50
枚の中になければ、今度は“君”の「ミ」か“きりぎ
りす”の「リ」か、2 音目で解はユニークにきまる。
即ち、“きまり字”は経過により刻々に変るから、STM
における経過の記憶も常に up date されていなければ
ならない。

われわれが、日常、会話をを行い、文章を綴る場合
も、言葉そのものは LTM からレトリブされるが、
各時点において、それまでの経過が STM になれば
文章にはならない。途中で話のコシを折られたりする
と、どこまで話していたかわからなくなることがある
のは誰しも経験するであろう。

3. LTM の検索

さきに述べたように、文章を綴る場合、何がキイに
なって必要な瞬間に必要な言葉が LTM から取り出さ
れているかは全く分っていないし、それに直結した実
験が行われたという話も知らない。しかし、次のよう
な実験なら容易に行える²⁾。一つのカテゴリーを指定
し、“それに属する単語をできるだけたくさん重複しな
いように想起せよ”という指令を与える。想起開始後
 t 分目までに被験者が書いた単語の累積個数 $n(t)$ を
 t に対しプロットしたのが図-2(次頁参照)である。同
一被験者が“漢字一字で書ける生物の名”(○印)、“子
でおわる婦人の名前”(×印)について想起を行った例で、

$$n(t) = n(\infty)(1 - e^{-\lambda t}) \quad (3.1)$$

という式がよくあてはまる。これは LTM の検索に関
する実験といえるであろう。より長時間にわたって単
語を想起させた例が図-3(次頁参照)に示してある。
このぐらいいく長くなると検索速度に小さな振動が現わ
れてくるが、基本経過としてはやはり (3.1) 式の形にな
っている。この想起過程について次の三つの点が指摘
できる。第一に、与えられたカテゴリー以外の単語が
思い出されることはなく、第二に、過程の後半になると
既に想起してしまった単語を再び思い出すことは珍

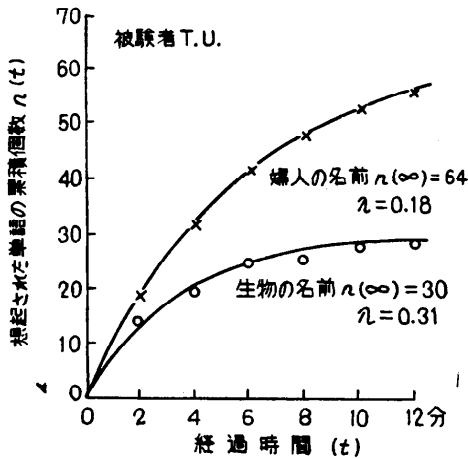


図-2 LTM からの retrieval の 1 例

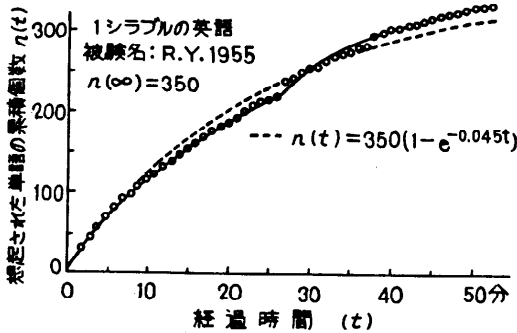


図-3 LTM からの retrieval (特に長い 1 例)

しくはないが、それを見落して同じ言葉を二度書くことはほとんどおこらない。第三は、関連のある単語がかたまって想起される傾向は認められるが、それは余り著しいものではない。この第三の傾向についてはここでは立入らず、第一、第二の点をめぐる考察だけを述べることにする。

カテゴリーが与えられると、LTM の中にある該当する単語には一勢に点燈（活性化）されるものとしよう。個々の単語がそれぞれ神経組織の局在された番地に記憶されているとは思えないが、機能的に一つの単位をなしていればそれでよい。第一の事実は該当しない単語が点燈されることがないという意味になる。ただし、後述するように、該当する単語の中で点燈されそこなうものはかなり出る。

上述の想起過程を確率過程と考えると、(3.1) 式は pure death process の期待値の経過に相当する。1 人の被験者についての 1 回の実験結果、即ち、何ら平均という操作を加えないでも、図-2, 3 のように安定し

たデータが得られるということは、この理由によるものであろう。この process においては、 $(n-1)$ 個目の単語と n 番目の単語の想起の間の時間間隔 $\tau(n)$ はいずれも指数分布に従い、相互に独立、且つ、そのパラメータ $\tau(n)$ の平均値の逆数 $\lambda(n)$ は n の線型減少関数

$$\lambda(n) = [n(\infty) - n(t)]\lambda \quad (3.2)$$

である。一方、 Δt を単位時間とし、 t と $(t + \Delta t)$ の間に生成される単語の期待数を $\Delta n(t)$ とおけば

$$\Delta n(t) = [N_p] - [N_p] \frac{n(t)}{n(\infty)} \quad (3.3)$$

$$\lambda = \frac{[N_p]}{n(\infty)} \quad (3.4)$$

と書ける。そこで、まず (3.4) 式について述べると、次のような検索過程を考えることが出来るであろう。この被験者は常に単位時間当り N 個の単語をチェックすることができ、その中に当該カテゴリーに属する単語（点燈されているもの）は常に P の確率で含まれている。ところが、ある単語がレトリブされても、それによって消燈されず、従って、次の検索の範囲から除外されないとすれば、想起が進むにつれ、 $[N_p]$ 個の中には既に想起された単語も含まれ、その混入の確率は $n(t)/n(\infty)$ である。以下、これを「定速・悉皆走査」モデルと呼ぶ。定速とは $[N_p]$ が t に依存しないこと、悉皆とは検出されても消燈されないことを指している。この型の検索が行われているとすれば、その後一つ一つのモニター機構があり、一旦想起された語はそこで排除されると考えなければ上述の第二の事実は解釈できない。これが (3.3) 式の右辺の第二項である。

一方、(3.2) 式から LTM のデータ・ベースについて次のようなイメージを描くことも可能である。LTM のエリアによって点燈した点の密度に大小があるものとし、検索は密度の高いところから徐々に密度の低いところへ移ってゆくものとすれば (3.2) 式は密度が n につれ直線状に減少することを示している。以下、これを「単語密度の線型減少」モデルと呼ぶ。この場合には第二の事実を説明するのに上述のモニター機構を想定する必要はないが、そのかわり、上述の形で走査エリアの変更を可能にするメカニズムの存在を想定しなければならない。

実験条件を少し変えると、想起過程 $n(t)$ の型も変わり、日本の都市を「北から南へ」思い出せというとき、図-4(次頁参照)になる。この場合には走査の道順がきめられているから、(3.3) 式でいえば右辺の第二項は不用になり、データ・ベースにおいて密度の変化も

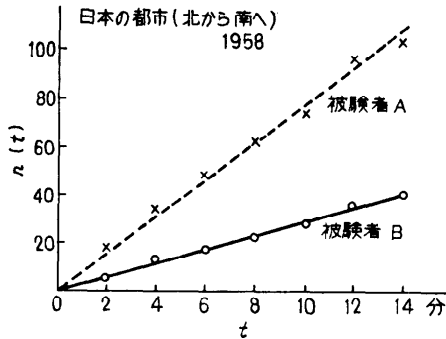


図-4 LTM からの retrieval (筋道のきまっている2例)

ない。また、同じカテゴリーについて、続けて想起をくり返すと $n(t)$ の関数型そのものが変化する。この事実がまた LTM の検索過程及び STM の介入について示唆を与えてくれるが、ここでは相当長期的の間隔において同じ人が同じカテゴリーについて想起を4回反復した場合を考える。この場合には $n(t)$ は常に(3.1)式の形になるが、漸近線 $n(\infty)$ だけが反復毎に少しずつその値を増し、かつ、 λ と $n(\infty)$ とはちょうど逆数関係になることが見出された(図-5)。この場合、人もカテゴリーも同じであるから $[N_p]$ を一定とすれば、これは(3.4)式に相当する。

4. STM の検索

この方は実験がやり易いので、古くから多くの実験が行われており、最近の実験結果はいずれも「直列定速悉皆走査」を示唆している。即ち、同一カテゴリーの n 個の単語のリストを暗記させ、直ちに一つのターゲットを示して、それがそのリストの中にあったか(Y)、なかったか(N)を判断するのに要する時間 $\tau_i(n)$

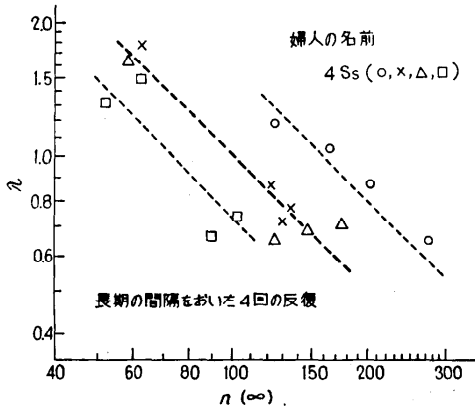


図-5 $n(\infty)$ と λ との逆数

を測定すると、 $\tau_Y(n)$ も $\tau_N(n)$ も同じ勾配 α をもって n の線型関数になる³⁾。ただし、切片 β の方は必ずしも一致せず、一致しない時には $\beta_N > \beta_Y$ の傾向がある。例えば、花の名前のリストに対してイヌというように、リストのカテゴリーとは全く別な単語をターゲットとして用いた場合 (N') には、STM を走査するまでもないので、 $\tau_{N'}(n)$ は n に依存しない定数に近くなり、 $\alpha_{N'} = 0$ である。 N' 以外の場合、 $\tau_i(n)$ が線型ということは直列定速を、 $\alpha_Y = \alpha_N$ ということは、Y の場合にも、N の場合と同じく、 n 個をすべてチェックしてからでない人間は反応しないことを示すものと考えられる⁴⁾。これがここでいう悉皆の意味である。もし、悉皆であれば、Y の場合、そのターゲットがリストの中でどの位置 k にあってたかによって、反応時間 $\tau(k)$ は変化しないはずであろう。たしかに $\tau(k)$ の平均値を位置 k に対してプロットすると、横軸に平行に近い結果となる(図-6)。ただし、悉皆以外の走査モードを考えても、この実験結果は解釈できないことはない。

一方、 n 個の単語を視覚的に提示、ターゲットを与えてその中から探させると、 $\tau_i(n)$ はやはり線型であるが、Y に対する α_Y は N に対する α_N の半分になる(図-7(次頁参照))。即ち、視覚検索においては、Y の場合、人間はターゲットを見出した瞬間に反応す

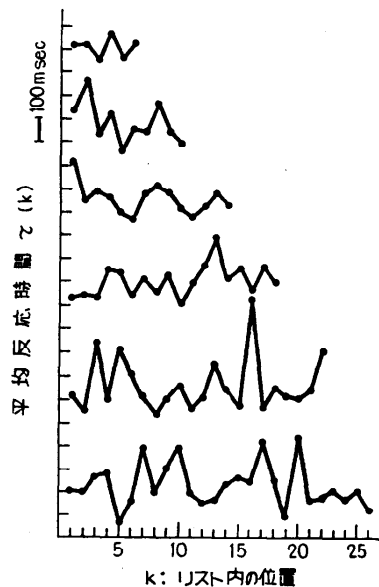


図-6 反応時間が記憶されたリスト内におけるターゲットの位置に依存しない事実

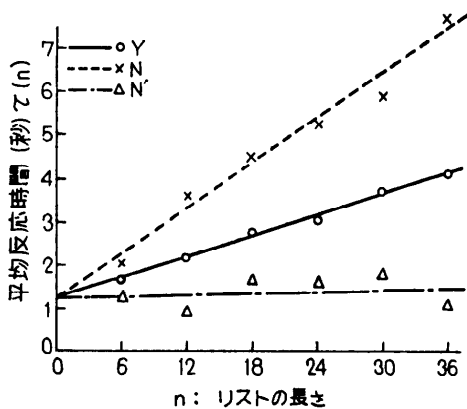


図-7 視覚走査(直列自動打ち切り型)

る「自動打ち切り走査」を行っているらしい。なお、勾配 α の値は、走査速度は視覚検索において大略 10 個/秒、STM の検索は、一度で記憶できる場合 ($n \leq 6$) には 30 個/秒、反復暗唱して記憶すると 100 個/秒のオーダーで、STM の検索は視覚検索に比べ著しく速い。視覚的素材は一種の外部記憶であるが、この点、人間もコンピュータと同じといえよう。百人一首において、視覚走査などやるのは素人で、選手級になると、専ら記憶に頼って札をとっている。上述の検索速度からいって、これは当然であろう。事実、一旦札の配列を眺めれば後は裏返しにされても困らないのである。

心理学の実験対象として考える限り、人間は一つのブラック・ボックスである。入力と出力とは外部から観測されるが、内部に作動しているソフトウェアは観察のしようがない。入力に相当するのは、3. の実験では所与のカテゴリー、本章の実験では記憶すべきリストやターゲットである。また、3. の実験では出力の時間の経過、本章の実験ではターゲットに対する応答とその潜時である。しかし、いかにして LTM の検索が行われたか、Y 或いは N の決定が行われたか、それを観察する手段はない。人脳をひらき電極を挿入しても、取り出されるのは神経組織の活動電位などの生体情報で、ここでいうソフトそのものではない。実は人間は第三者にとってブラック・ボックスであるのみならず、本人自身にとってもブラック・ボックスなのである。3. の実験において、“どうやって単語を探しましたか”と尋ねても、“一生懸命探しました”というだけで、LTM の IR やデータ・ベースに関する何の情報も与えてくれない。本章の STM の走査に関する実験においても、私自身が被検者として、Y の場合、

悉皆走査などバカなことをやらないつもりで応答しても、やはり $\alpha_Y \equiv \alpha_N$ という情けない結果に終るのである。

5. 直接的 pop up

LTM からの想起において、定速悉皆走査モデルの (3.3) 式の形は、時点 t において、取り上げられた個々の単語が既にレトリブされたものか否かに関し STM をチェックするのに要する時間は常に無視得るという意味を含むであろう。上述のように、STM の走査は極めて高速という点から、これは首肯できるが、更にそれが $n(t)$ に依存しないということについて、最近、筆者と金沢は次の実験事実を示すことができた⁵⁾。実験者から与えられて記憶したリストの場合には $\tau_i(n)$ はたしかにリストの長さ n の線型増加関数になるが、想起の場合、被験者が n 個をレトリブしたところで、一旦、想起をとめてターゲットを与えると、 $\tau_i(n)$ はほとんど n によって変化しない。ただし、想起とターゲットの提示との間に数日をおくと、 $\tau_i(n)$ は n につれて増加する形にもどってくるのである。与えられたリストを暗記して STM につめ込んだ情報と、自ら生成したリスト、いわば、ひとりでに STM も入った情報との間にはその検索に関し、著しく異なるものがあるらしい。われわれが会話をを行い、文章を綴っている時にも、経過が STM に入っていないければならないことは 2. の末尾に指摘した通りであるが、これも自ら生成したリストである。一般に、STM におけるこの種の情報の検索に走査という操作が含まれているか否か、これはすこぶる疑わしい。特に特定のカテゴリーに属する単語をすべてダンプさせる 3. の実験の場合、既に想起された単語が LTM から再び取り出され、それが既出のものであることが検出されて排除されたとしても、それが STM の検索の結果であるか、或いは、単語自身に既出というタグがつけられていて、取り出された瞬間にそれで検出されているのか、現在のところよく分らないのである。

今までの話の中に、カテゴリーがきめられると、LTM 中の該当する単語が一勢に点燈するという仮設があった。人間が会話をを行い、文章を綴るに当たっても、該当する言葉は、一語一語、直接に pop up してくるようで、LTM を走査するという操作が含まれているか否か、すこぶる疑わしい。ここに、人間の検索とコンピュータの検索とが著しく異なる点があるようで、人間の場合、必要な情報が一個だけ、或いは、そ

のすべてが同時に pop up するという現象をどうしても考えたくなる。会話を可能にしているメカニズムの中にも同じ機構が含まれているであろう。この直接的、自発的 pop up の場合には、走査によりキイにマッチするものを見出すというより、共鳴に近い現象が作用しているのであるが、そのメカニズムについて現在は何の手がかりも持っていない。ただし、キイの中にこの共鳴を可能にするものとそうでないものがあるらしく、単語の意味内容、自然に存在するカテゴリー、単語の第一音などをキイとする場合、該当する単語が直接的に pop up し、それ以外の単語を思いつくことはまずないのに反し、第二音をキイとし、2 番目が「ア」の単語を探そうとすれば、幾つかの単語を、一旦、ズラズラ思い浮かべてからでないと拾えない。どの辺を探せば該当する単語がありそうかという漠然とした予感があり、それに従って検索範囲をせばめているにしても、検索は間接的な形になっている。

図-8 の(A)を見ると、自発的に一对のマル印が浮かび上って認知されるであろう。一方、(B)になると、客観的には同じマルの対が存在しているのであるが、それは探す気にならなければ認知できないであろう。共鳴的な作用と走査による結果との相違である。STM の検索においても同じことがおこる。

$$21\left(\frac{91}{7}+6\right)-14=$$

の答を求めるには、 21×19 を計算する必要があるが、これは $(a+b)(a-b)=a^2-b^2$ を用いると、暗算で簡単に $400-1=399$ となる。普通の人達にこの便法を教えながら、一定の時間、(A)グループにはマッチで図形をつくるという計算とは無関係の作業をやらせ、(B)

グループにはひき続き計算問題をやらせる。ただし、この計算には上述の便法を用いる余地はない。ついで、両グループの人々に

$$(15+64-47)28+(-20+34)=$$

の答を求めさせる。この中には 32×28 という計算が含まれており、便法を用いると $900-4=896$ と簡単である。ここで自発的に便法を思い出し、それを用いて計算した人の率を見ると、(A)グループでは 0.73、(B)グループでは 0.26 であったと報告されている⁶⁾。自発的に使用しなかった人も、尋ねて見れば、それを忘れてしまっていた訳ではない。共鳴的作用が生じし易いか否かは、図-8 と同じように、当該ペアーを囲んでいる周囲の条件に依存するのである。STM からにせよ、LTM からにせよ、適当なキイが与えられれば、キイが一方のマル、該当する情報が他方のマルとして作用し、走査を経ないでそれを pop up させる content addressing ともいうべき何かがあるのであろう。これをコンピュータにおける associative memory と対比させてよいものか否かは筆者にはわからない。

6. 推理能力の測定

記憶や検索よりも高次な情報処理過程、例えば推理などになると、その実験的説明は更に困難になる。しかし、推理、問題解決の能力に大きな個人差が存在することは事実で、当学会などは最もそういう能力に恵まれた方々の集まりであろう。そこで、個人々々の推理能力の優劣を測定の名に値するような形で表わす心理テストを構成してみた。これを LIS 推理因子測定尺度と称する。LIS というのは F.M. Lord の理論にもとづき印東、鮫島が作ったところから名づけた^{8),9)}。

一般に、心理テストといわれるものの多くは、定められた時間内にどれだけ問題をこなせるかという処理速度の大小でスコアがきまるが、これを speed limit という。しかし、今回は、いずれもじっくり考えて解く問題、十分時間を与えても解けない人には解けないというタイプの問題を $n=30$ 題用意した。これを power limit という。この 30 題のすべてを行うにはゆうに 4 時間を要する。個々の問題をアイテムと呼ぶが、各アイテム j に対する応答は正答か否かで 1 或いは 0 の値をとる二値変数 X_j で、これは観測できる。人間の母集団は東京及びその近郊の中学生 2, 3 年生 (1960 年当時) と定義してあるが、 $N=883$ 名のサンプルについて X_j を得た。アイテム j, k に対する応

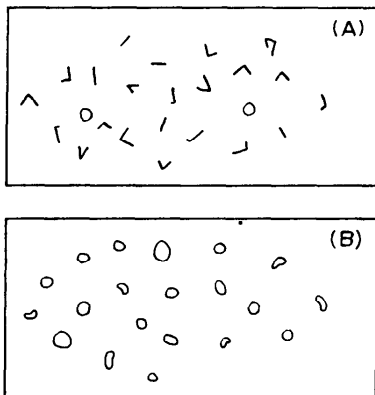


図-8 自発的な一对形成と検索により一对の発見

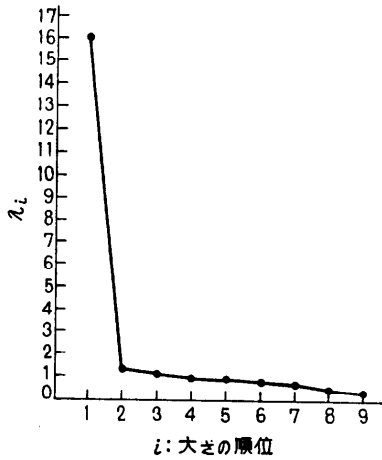


図-9 推測因子の存在 (30 題の推測問題の相関行列の固有値)

答 X_j, X_k を組み合わせると、 2×2 の表ができ、四分位 (tetrachoric) 相関係数 r_{jk} がきまり ($-1 \leq r_{jk} \leq 1$)、 $n \times n$ の行列 R を得た。この R の固有値 λ を計算すると図-9 になり、一つだけ (λ_1) が大きな値をとり、他の λ_i はいずれもゼロに近いレベルにある。この事実は、 $n=30$ 題のアイテムに対し、それに回答する心理過程の底にただ一種類の心理過程の存在を想定してもよいという意味をもち、以下、その優劣を連続量 c の値で表わすことになる。 λ_1 に対応する R の固有ベクトルとアイテム j の正答率とを利用すると、能力がある値 c の人達が問題 j に正答し得る条件つき確率 $P_j(c)$ がきまり、これをアイテム特性曲線という。

テストの得点 s を正答できたアイテムの数によって表わすとすれば、

$$s = \sum_{j=1}^n X_j \quad (6.1)$$

で、能力のレベルがある値 c の人がある得点 s をとる条件つき確率 $P(s|c)$ は

$$P(s|c) = \sum_s \prod_{j=1}^n P_j(c) X_j Q_j(c) 1 - X_j \quad (6.2)$$

$$Q_j(c) = 1 - P_j(c)$$

$\sum_s$ は (6.1) 式をみたす X_j のすべての組み合わせについて和をとることを表わす。

となる。(6.2) 式は

$$\prod_{j=1}^n (P_j(c) + Q_j(c)) \quad (6.3)$$

の展開項に当るので、これを拡張された二項分布とい

い、その平均と標準偏差は次の形をとる。

$$\bar{s}(c) = \sum_{j=1}^n P_j(c) \quad (6.4)$$

$$\sigma_s(c) = \left[\sum_{j=1}^n P_j(c) Q_j(c) \right]^{1/2} \quad (6.5)$$

$\bar{s}(c)$ をテスト特性曲線というが、 c の全範囲にわたって $\bar{s}(c)$ が線型にはなり得ないことは容易に示される。

$$\hat{s}(c) = \lim_{n \rightarrow \infty} \frac{\bar{s}(c)}{n} \quad (6.6)$$

とおけば、 $\hat{s}(c)$ のまわりの s の分布の標準偏差は 0 になり、アイテムの数を大きくしてゆけば、ある値 c の人の得点は 1 点 $\hat{s}(c)$ に集中してくるのである。(6.2) 式の根底には上述の λ_1 だけが正の値をとること、即ち、 X_j, X_k を結びつけている ($r_{jk} > 0$) のは c の存在だけで、 c を固定すればアイテム j, k に対する回答は独立になるという想定が効いている。次に、人間の母集団における c の密度関数 $f(c)$ さえきまれば、 c と s との 2 次元密度関数は

$$f(s, c) = \int_{-\infty}^{\infty} P(s|c) f(c) dc \quad (6.7)$$

で計算できる。ここでは $f(c)$ は正規分布としておく。本来、変数 c については、それが存在するという保証だけしか与えられていないので、未だそのスケールは定義されていない。従って、以下、 $f(c)$ が正規分布になるよう c のスケールを約束したと思って頂きたい。こうして、 $f(s, c)$ がきまれば、その周辺度数が得点 s の理論的度数分布 $f(s)$ であるから、 $f(c)$ と実測された得点 s の度数分布 $f(s)$ とを比較することができ、それが図-10 である (LIS-P 版)。

その特性曲線 $P_j(c)$ が既知のアイテムが十分にあるものとし(アイテム・バンク)、その中から、適当な n

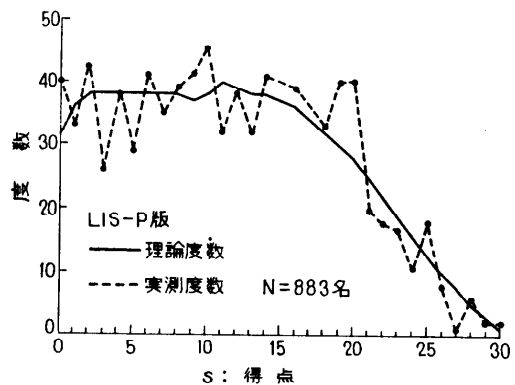


図-10 LIS-P 版におけるスコアの理論度数分布と実測度数分布

題を組み合わせでテストを構成すれば、そのテストはどのような得点の分布 $f(s)$ になるか予測できる。従って、目的に応じ、適当な題数 n をもって都合のよい $f(s)$ をもつテストをつくるのが可能になる。例えば、プログラマーとして優秀な人だけを1時間半位で選抜したいとしよう。 c の値の低いところは問題にならず、 c のレベルの高い人々についてのみ、得点なるべく同点に重ならないようにすればよいから、なるべく右に尾をひいた分布 $f(s)$ をもつテストが望ましい。上述の30題の中からある $n=7$ 題を選んで一つのテストをつくると、 $f(s)$ は実際にそういう型になる (LIS-S版)。 $f(s, c)$ が与えられているのであるから、信頼係数 $\beta=0.98$ をもって、得点 s から c の値の区間推定を行うことができる。図-11の横線は、得点 s をとった $i=1\sim 100$ 人中98人までは、 c_i はその中にあると推定される区間を示し、測定の精度に当る。もちろん、30題の場合 (LIS-P版) と上述の7題 (LIS-S版) とでは区間の大きさは異なる。現在のところ、P版ぐらいが実際に到達できる心理測定の限界ではないかと思われる。

この程度になると、これは心理テスト、検査というよりは、心理測定というべきであろう。もちろん、測定の精度に関しては計測器による物理測定と比肩でき

ないまでも、その論理構造において、少なくとも初期の温度測定と大差はない。直接には見えないながら、スカラー c で表わされる測定対象の存在が保証され、それと観測結果 s との対応及びその精度も明示でき、原理的には所定の対応、精度をもつよう心理測定を設計することもできるからである。ただし、 c のスケールには絶対的な意味は与えられない。先述のように、所与の母集団において密度関数 $f(c)$ がきめられた平均 (スケールの原点の定義)、標準偏差 (スケールの単位の定義) をもつ正規分布になるスケールとしか意味がつけられない。温度の測定でも、1660年頃フローレンスで作られた初期の温度計には“もっとも厳しい冬の寒さ”とか“バターが融ける夏の温度”などという原点、単位でスケールが刻まれていたそうで、こういう計測による研究を蓄積した後でなければ、絶対温度という概念には到達できない。心理測定も、現段階はルネサンスにおける温度測定のレベルにあるといつてよく、計測のスケールも任意の約束以外に方法はなく、計測されている現象の本体もさだかではないが、約束だけははっきりした計測とそれによる観測とを地道に続けてゆく他ないであろう。

7. あとがき

過日、The Psychology of Computer Programming (1971, Van Nostrand) の著者である G. M. Weinberg 教授は、現在のソフトウェアは第一義的にはハードの制約の下にあること、こうして構造のきまったソフトに人間が無理して歩みよって用いていること、将来はもつとソフトを人間に近づけるべきことを強調された (第9回 IBM コンピュータ・サイエンス・シンポジウム)。人間側の制約にあわせてソフトの構造を考えるとすれば、人間の情報処理機構に関し、より豊富な数量的データを整備しておく必要があるとし、能力の個人差を数量的に表現するより精緻な方法も必要であろう。本稿に述べたのは、そういう指向なしに行われた研究に過ぎないが、例えば大規模ソフトウェア開発の実技の中から適当な心理学的実験、測定のテーマを指示して頂ければ、より直接に役立つデータを蓄積してゆくことも不可能ではないであろう。人間についての研究は遅々として進まないから、基礎研究の中から有用な情報を得ようとしていたのでは、いつになっても応用は効かない。ここに述べたようなスタイルのものでなければ、問題さえ与えれば、それと取り組むことはできるであろう。人間の研究が進まないこと

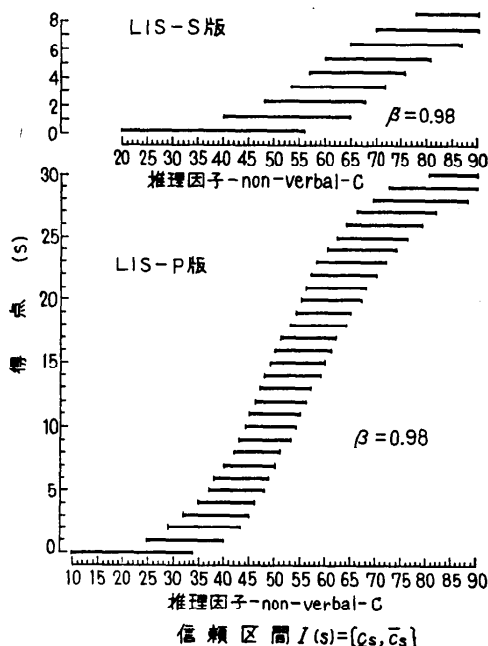


図-11 スコア s による因子 c の区間推定 (LIS-P 版及び S 版)

には幾つかの理由がある。第一に、人間について破壊試験をやる訳にはゆかない。第二に、観測可能な量も限られていて、本当に頼れるのは時間の計測と計数ぐらいなものである。本稿においても、テストの得点 s や、時間上の個数 $n(t)$ という計数の観測を基礎にしていた。第三に、人間に関する観測結果は、多くの場合、context free ではないから、ある状況の下で得られた結果が他の場合においてもそのまま成立するとは必ずしも保証できない。従って、本当に役立つデータを必要とすれば、やはり、それが必要とされる場合に近い状況の下でそれを獲得する他はないのである。

参 考 文 献

- 1) 印東太郎, 久野 麗: ピアノを用いて行った把持曲線に関する解析的研究, 哲学 (慶応義塾大学), 第 32 輯, pp. 155~172 (1956).
- 2) T. Indow & K. Togano: On retrieving sequence from long-term memory, *Psychol. Rev.*, Vol. 77, No. 4, pp. 39~43 (1970).
- 3) S. Sternberg: High-speed scanning in human memory, *Science*, 153, pp. 652~654 (1966).
- 4) T. Indow & A. Murase: Experiments on memory and visual scannings, *Jap. Psychol. Res.* Vol. 15, No. 3, pp. 136~146 (1973).
- 5) T. Indow & H. Kanazawa: Recognition memory of retrieved sequence of words, (in printing).
- 6) W. Kohler & H. von Restorff: Zur Theorie der Reproduktion, *Psychol. Forsch.*, Vol. 21, pp. 56~112 (1935).
- 7) F.M. Lord: A theory of test scores. *Psychometric Monographs*, No. 7 (1952).
- 8) 印東太郎, 鮫島史子: LIS 推定因子測定法——non-verbal——日本文化科学社, (1962).
- 9) T. Indow, & F. Samejima: On the results obtained by the absolute scaling model and the Lord model in the field of intelligence. 3rd Report, The Psychological Laboratory on the Hiyoshi Campus, Keio University, (1966).