



## 多段階教授システムの最適化\*

溝口 文雄\*\*

### Abstract

The present study is carried out from standpoints of multi-stage decision problem to determine, through dynamic programming, the optimal instructional items and sequences. We also considered the various factors which affected the learner's assessment in the selection of offered courses which had different type of criteria.

We discussed the merits and restrictions of the approach based on the stochastic learning theory in the synthesis of optimal instructional system through the actual case studies. Finally, we further propose that an instructional system must be equipped with the structure of criterion which reflects the costs of information processing for the instructional feedback, i. e., diagnosis for the start of learner's understanding.

### 1. はじめに

教授システムの目的は学習者にとって最適な教授情報を与えることにある。とくに、CAI (computer aided instruction) の場合、この目的を実現するためには次の2つの理論的側面から検討する必要がある。第1は、学習者の学習のプロセスをどのような形で記述できるか、第2は、学習者の学習のプロセスをどのようにして制御し、最小の教授干渉で最適な教授活動ができるかである。通常、前者は学習の理論、後者は教授の理論と言われている (Brunner, J. S., 1964<sup>9)</sup>。

学習の理論のうち、もっともよく知られているのは行動主義の中心である刺激-強化 (S-R) 理論である。これに対して、記憶の構造を基礎におく、認知理論も最近になって発展し、人間の情報処理システムという視点から学習の理論を構築している (Neisser, U. 1967)。したがって、教授システムの構成は、どのような学習の理論を採用したかによって教授政策が異なってくる。もし、学習者の状態を刺激-反応理論または確率的学習モデルにもとづいて教授最適化を実現しようとすると、最適化は最小のエラー確率で学習するための

教材項目の提示順序を動的計画法を用いて決定するという多段階最適化問題に帰着される。たとえば、Groen, G. J., Atkinson, R. C.<sup>5)</sup> (1966), Atkinson, R. C., Paulson, J. A. (1972)<sup>1)</sup>, Matheson, J. (1964)<sup>8)</sup>, Smallwood, R. D.<sup>2)</sup> (1971)<sup>2)</sup> らの研究が、この多段階最適化のアプローチである。これに対して、認知理論では、学習者のもつさまざまな仮設とその検証、そのための情報の検索といった面に注目している。そして、教授干渉に、学習者自身のそうした意図的活動をどの程度まで取り入れれば最適化が実現できるかという、いわば、教授システムの制御能力と学習者に与えられた学習上の意志決定、例えば教授コースが自分に適しているか、のペイオフ問題と考えられる。すなわち、もし、教授システムが学習者の主観的働きを無視すれば、学習者は全く受動的に学習していくことになり、逆に、すべてを学習者の意図的働きにまかせたときには、教授システムの干渉能力は無くなり、システムの効率は失われる。したがって、両者のバランスを保つことによってなされる教授最適化のためのペイオフが存在することになる。こうした点を考慮した方向として Kopstein & Seidel<sup>7)</sup> (1970) の IMPACT、また溝口らの研究<sup>12)-14)</sup>がある。

本研究は立場によって異なる教授最適化の概念を明確にし、お互いにどこまでが有効であり、限界がある

\* Synthesis of the Optimal Multi-stage Instructional System by Fumio MIZOGUCHI (Department of Industrial Administration, Science University of Tokyo)

\*\* 東京理科大学理工学部経営工学科

かを検討する。そのために確率的学習モデルを前提とする多段階の教授プロセスについては、動的計画法により定式化を行う。次に、具体的例題にこの定式化を適用し、多段階の教授プロセスの最適化を検討する。最後に、教授システムと学習者とのペイオフの問題と教授システムが行う学習者への理解状態の診断の問題を取り上げ、教授最適化に必要な認知的側面の有効性を検討する。

## 2. 多段階教授システムの理論

### 2.1 適用領域と数学的定式化

多段階教授システムでは教授プロセスを離散的マルコフプロセスと考える。そのときシステムの状態は有限の状態であり、次の状態に移る時間は一定である。しかも、現在の状態  $i$  のとき、次の時間区間の間に状態  $j$  へ移る確率は  $i$  と  $j$  だけの関数によるものとする。このように非常に限定した教授システムが実際の教授プロセスのうち適用できる分野として考えられるのは、教材が文脈に依存しないような CF ドリル・訓練型 (context free drill and practice) である。この場合、学習者の状態は教材の情報構造とは無関係に、ドリルの回数に依存して記述される。例えば、Atkinson et al. の教授システムのように、英単語のスペリングのドリルでは、学習状態は以前に提示された回数、あるいはスペルを記憶するため学習者が行ったリハーサル回数に依存していると考えられる。このときは、学習状態は教材の提示経過に伴って起こる推移だけを考えればよい離散状態として扱える。すなわち、CF ドリル・訓練型の教授プロセスを離散的マルコフと考えていけばよい。あるいは、このように限定することにより、CF ドリル・訓練型の教授プロセスの定式化が行え、教授コースの生成順序の設計が可能になってくる。実際の有効性については、本章の 2.3 で取扱う例題での実験から検討する。そして、多段階教授システムとしたときの最適化が教授システム構成のうちのどのような面に有効であったかを、2.3.2 でも議論する。学習状態をマルコフプロセスで記述できないような学習者のもつ、この程度なら理解でき、その結果得るスコアはこれくらいというようなペイオフの問題は次節の 2.4 で考えていく。

まず、多段階教授システムを次のように定義する。

i)  $S$  をシステムの状態の集合とする。

$$S = \{s_1, s_2, \dots, s_N\}$$

ただし、 $N$  は全体の段階数で、 $N$  は  $N$  段階の

プロセスよりなるものとする。

ii)  $D$  をシステムによってなされた教授決定の集合とする。

$$D = \{d_1, d_2, \dots, d_N\}$$

iii)  $f$  を関数:  $S \times D \rightarrow S$  とし、あと  $n$  段階残された  $s_n, d_n$  に対して次の関係より新しいシステムの状態  $s_{n-1}$  を逆むきに決定していく。

$$s_{n-1} = f(s_n, d_n) \quad (1)$$

iv)  $d$  を教授上の決定政策とし、あと  $n$  段階残されたときに  $D$  の集合をシステムの各々の状態に、次の関係にしたがって割当てていくものとする。

$$d_n = d(n, s_n) \quad (2)$$

v)  $g$  を学習者に与える評価関数、例えば、各段階で学習者の得るテストスコアとする。  $n$  段階での評価関数  $g_n$  はシステムの状態  $s_n$  と教授決定  $d_n$  によって次の関係から決められる。

$$g_n = g(s_n, d_n) \quad (3)$$

最終段階  $N$  を次のように定義する。

$$g_N = G(s_0) \quad (4)$$

$s_0$  を最終段階のシステムの状態とする。

vi) 教授決定  $d_n$  のもとで、システムの状態が  $s_n$  より  $s_{n-1}$  に推移する確率を  $P(s_{n-1}|s_n, d_n)$  とすると、その時の  $g_n$  の期待値 ( $E$  と書く) は次の関係より求められる。

$$\begin{aligned} E[g_n] &= E[g(s_{n-1}, s_n, d_n)] \\ &= \sum_{s_{n-1} \in S} g(s_{n-1}, s_n, d_n) \cdot P(s_{n-1}|s_n, d_n) \end{aligned} \quad (5)$$

以上のように多段階の教授システムを定義したとき、問題は次のように記述できる。

問題: 教授システム ( $I_s$ ) を決定政策  $d$  のもとでシステムの状態  $s_n$  より  $s_{n-1}$  になるときの評価関数の期待値を最適にせよ。

$$I_s: \text{Optimal} \Leftrightarrow \text{Max} \left\{ E \left[ \sum_{n=N}^1 g(s_{n-1}, s_n, d_n) + G(s_0) \mid d \right] \right\} \quad (6)$$

すなわち、問題は各段階での期待値の総和を決定政策  $d$  のもとで最大にすることに帰着される。最適な決定政策を  $d^*$  と書くと、 $d^*$  はシステムの状態  $s_n$  より  $s_{n-1}$  に、次の条件を満足する決定政策である。

$$\begin{aligned} E \left[ \sum_{n=N}^1 g(s_{n-1}, s_n, d_n) + G(s_0) \mid d^* \right] \\ - E \left[ \sum_{n=N}^1 g(s_{n-1}, s_n, d_n) + G(s_0) \mid d \right] \geq 0 \end{aligned} \quad (7)$$

多段階の教授過程の各段と最終段の評価の和の期待

値を期待評価値として  $W_{d_n}(n, s_n)$  と書くと、あと  $n$  段階残された状態  $s_n$  にあるときの決定  $d_n$  のもつ期待値を意味している。次に、 $n$  段階残されたときの各段階が最適決定がなされたときの期待値を最適期待評価値として  $W(n, s_n)$  とすると、次の関係になる。

$$W(n, s_n) = \max_{d_n} W_{d_n}(n, s_n) \quad (8)$$

最終段階では、

$$W(0, s_0) = G(s_0) \quad (9)$$

となる。

残り 1 段階の期待評価値は次のようになる。

$$\begin{aligned} W_{d_1}(1, s_1) &= E[g(s_0, s_1, d_1) + G(s_0 | s_1, d_1)] \\ &= \sum_{s_0 \in S} [g(s_0, s_1, d_1) + G(s_0)] P(s_0 | s_1, d_1) \\ &= \sum_{s_0 \in S} g(s_0, s_1, d_1) P(s_0 | s_1, d_1) \\ &\quad + \sum_{s_0 \in S} G(s_0) \cdot P(s_0 | s_1, d_1) \end{aligned} \quad (10)$$

(5)式を用いて

$$\begin{aligned} E[g_1] &= E[g(s_0, s_1, d_1)] \\ &= \sum_{s_0 \in S} g(s_0, s_1, d_1) \cdot p(s_0 | s_1, d_1) \end{aligned} \quad (11)$$

および(8)式から

$$\begin{aligned} W(1, s_1) &= \max_{d_1} \left\{ E[g(s_0, s_1, d_1)] \right. \\ &\quad \left. + \sum_{s_0 \in S} W(0, s_0) \cdot P(s_0 | s_1, d_1) \right\} \end{aligned} \quad (12)$$

最適決定政策を  $d^*$  により、次式が成り立つときに(12)式は最大となる。

$$d_1 = d^*(1, s_1) \quad (13)$$

一般の  $n$  段階についても同様に導くことができる。

すなわち、最適決定政策は再帰的に計算することができるので、 $n$  よりも小さな  $k$  の最適決定政策  $d^*(k, s_k)$  から  $d^*(n, s_n)$  が算出される。

したがって、

$$\begin{aligned} W_{d_n}(n, s_n) &= E \left[ g(s_{n-1}, s_n, d_n) \right. \\ &\quad \left. + \sum_{k=n-1}^1 g(s_{k-1}, s_k, d^*(k, s_k)) \right. \\ &\quad \left. + G(s_0) | s_n, d_n \right] \end{aligned} \quad (14)$$

$$\begin{aligned} &= E[g(s_{n-1}, s_n, d_n) | s_n, d_n] \\ &\quad + E \left[ \sum_{k=n-1}^1 g(s_{k-1}, s_k, d^*(k, s_k)) \right. \\ &\quad \left. + G(s_0) | s_n, d_n \right] \end{aligned} \quad (15)$$

条件付期待値の性質を考えると、

$$\begin{aligned} &E \left[ \sum_{k=n-1}^1 g(s_{k-1}, s_k, d^*(k, s_k)) + G(s_0) | s_n, d_n \right] \\ &= E \left[ E \left[ \sum_{k=n-1}^1 g(s_{k-1}, s_k, d^*(k, s_k)) \right. \right. \\ &\quad \left. \left. + G(s_0) | s_{n-1}, s_n, d_n \right] | s_n, d_n \right] \end{aligned} \quad (16)$$

内側の期待値は  $s_{n-1}$  の条件のもとで生起する事象にのみ依存しているため、 $s_n, d_n$  とは独立している。

したがって、条件  $s_{n-1}, s_n, d_n$  はマルコフプロセスの性質により条件  $s_{n-1}$  にすることができる。この性質を(16)式に適用し、 $W_{n-1}(n-1, s_{n-1})$  の項を用いて次式を得る。

$$\begin{aligned} &E \left[ \sum_{k=n-1}^1 g(s_{k-1}, s_k, d^*(k, s_k)) + G(s_0) | s_{n-1}, s_n, d_n \right] \\ &= E \left[ \sum_{k=n-1}^1 g(s_{k-1}, s_k, d^*(k, s_k)) + G(s_0) | s_{n-1} \right] \\ &= \sum_{s_{n-1} \in S} W(n-1, s_{n-1}) \cdot P(s_{n-1} | s_n, d_n) \end{aligned} \quad (17)$$

式(15)、(16)、(17)より、整理して次式を得る。

$$\begin{aligned} W_{d_n}(n, s_n) &= E[g(s_{n-1}, s_n, d_n) | s_n, d_n] \\ &\quad + \sum_{s_{n-1} \in S} W(n-1, s_{n-1}) \cdot P(s_{n-1} | s_n, d_n) \end{aligned} \quad (18)$$

これより、一般式として最適期待評価値は次式で表わされる。

$$\begin{aligned} W(n; s_n) &= \max_{d_n} \left\{ E[g(s_{n-1}, s_n, d_n) | s_n, d_n] \right. \\ &\quad \left. + \sum_{s_{n-1} \in S} W(n-1, s_{n-1}) \cdot P(s_{n-1} | s_n, d_n) \right\} \end{aligned} \quad (19)$$

式(19)は次の決定政策の時に最大値をとる。

$$d_n = d^*(n, s_n) \quad (20)$$

以上の多段階教授システムの基礎方程式は(19)式で、その最適解の条件は(20)式を満たす  $d_n$  を見い出していけばよい。

## 2.2 多段階教授システムにおける評価の検討

今までに述べた多段階教授過程の定式化では、各段階の期待評価値を考えた。しかし、各段階への推移が学習者の学習プロセスのモデルから記述できる時には、教授システムの状態はツリーダイアグラムで確定論的に記述できる。したがって、最適決定の教授項目の順序は各段階の評価関数  $\{g_n \rightarrow \text{Max. } g(s_n, d_n)\}$  となるツリーを選んていけばよい。すなわち、次の形式で表現される。

教授システム  $I_s$ ,

$$\text{Optimal} \Leftrightarrow \text{Max} \left\{ \sum_{n=N}^1 g(s_n, d_n) + G(s_0) | d \right\} \quad (21)$$

評価関数の期待値を最適にする時には、システムの状態  $s_n$  から新しい状態に逆むきになる  $s_{n-1}$  への確率を考慮する必要があった。(21)式ではその項だけ簡略化される。システムの状態の変化は学習者の学習プロセスに対応している。確定的に学習プロセスが記述できる時の評価は式で表現されるので、問題は学習プロセスをどのように記述していくかになる。通常、学習者の学習状態を記述するもっとも簡単なモデルは単一オペレータモデル (Single operator Model) と単一要素モデル (one element Model)<sup>1)</sup> である。

Atkinson & Paulson の言語教授システム<sup>1)</sup>、およびそれに関連する藤田らの CAI タイプトレーナー<sup>4)</sup> (1972)の研究でも、上記の学習モデルを評価の基本にしている。

評価関数の期待値による教授プロセスとの比較の意味から、本研究では具体的方法として、まず上記の学習モデルを想定する単語教授システムをミニコンピュータ FACOM U-200 で実現し、実験的にシステムの有限性と限界を調べる。次に各段階への評価を期待値として扱う例に、教授システムは選択に際して学習者のもつベイオフを取りあげ、シミュレーションより教授プロセスを検討し、多段階教授プロセスの評価問題を考察する。

### 2.3 例題 1 (Atkinson タイプの教授最適化)

評価の基本になる学習モデルを吟味することから述べる。単一要素モデルでは、学習者は教授項目に対して学習状態  $L$  と未学習状態  $L^0$  とのいずれかひとつの状態にある。教授項目が状態  $L$  にあるときは、学習者は項目に対して正しく応答する。それに対して、項目が状態  $L^0$  にあるときは、確率  $q$  で正しく応答する。教授項目が各段階の始めに、状態  $L^0$  にあり、これが提示されたとき確率  $c$  で状態  $L$  に変わっていく (ただし、 $0 < c < 1$ )。これを簡単な状態グラフで表現すると Fig. 1 のようになる。そこで、次のようにして学習者の状態を記述する。

$$s_n^{(j)} = \begin{cases} 1 & \text{項目 } j \text{ の } n \text{ 回目の提示で誤応答} \\ 0 & \text{項目 } j \text{ の } n \text{ 回目の提示で正応答} \end{cases}$$

$P_r$  を確率とすると、 $s_n^{(j)} = 1$  になる確率、および  $(n-1)$  回目が  $s_{n-1}^{(j)} = 1$  のときに、 $n$  回目に  $s_n^{(j)} = 1$  となる確率はそれぞれ次式で表わされる。

$$P_r(s_n^{(j)} = 1) = (1-q)(1-c)^{n-1} \quad (22)$$

$$P_r(s_n^{(j)} = 1 | s_{n-1}^{(j)} = 1) = (1-q)(1-c) \quad (23)$$

これに対して、単一オペレータモデルでは、学習者は特定の教授項目  $j$  に対して、 $s_n^{(j)}$  を  $n$  段階にお

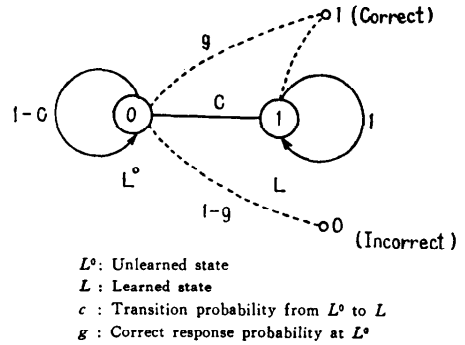


Fig. 1 The all or none model

るエラー確率として、次の状態  $s_{n+1}$  に次式で変わる。

$$s_{n+1}^{(j)} = s_n^{(j)} \cdot \alpha \quad (24)$$

ただし、 $0 < \alpha \leq 1$  とする。

または、変形して

$$s_{n+1}^{(j)} = \alpha^n \cdot s_1^{(j)} \quad (25)$$

となる。これは単一要素モデルにおいて、 $\alpha = 1 - c$ 。

$s_1 = 1 - g$  のときの  $P_r(s_n^{(j)} = 1)$  の表現は次のようになる。

$$P_r(s_n^{(j)} = 1) = \alpha^{n-1} \cdot s_1^{(j)} \quad (26)$$

したがって、単一要素モデルの(23)式を考えれば、単一要素モデルも、単一オペレータモデルと同じ表現になる。ここでは、簡単のために単一オペレータモデル(25)式を用いて多段階は教授プロセスを最適化する。例題を次のように設定する。

『 $K$ 個の教授項目を、 $N$ 段階で教授するとき、その  $K$ 個の教授項目を  $S$ 個よりなる  $j$ 個サブ項目に分割し、最適な提示順序を決定せよ。』

学習者の状態  $S$  は  $i$ 段階において  $m$ -項目と定義する。

$$S = \{S^{(1)}, S^{(2)}, \dots, S^{(m)}\}$$

ただし、 $S^{(j)}$  を項目  $j$  に対するエラー確率で、(25)式に従うものとする。

評価関数は教授プロセスの最終段における正応答の確率を最大にすることとすると、次のようになる。

$$W(n, s_n) = \text{Max} \sum_{j=1}^m \{1 - s_n^{(j)}\} \quad (27)$$

ここで、 $K$ 個の教授項目が  $S$ 個よりなる  $J$ 個のサブ項目に分割されているので、教授項目の提示順序は  $j$ 個の教授決定からなされる。すなわち、次の集合になる。

$$D = \{d_1, d_2, \dots, d_j\}$$

したがって、最適政策は、教授項目  $j$ 個のなかから、(27)式を最大にする決定を  $D$ より選ぶことになる。

最適政策は次のステップから求められる。

**Step 1.** 第  $N$  段階のプロセスを第 1 段階と考えると学習者のすべて可能な状態を初期のエラー確率とともに (25) 式より計算する。

**Step 2.** 第  $N$  および第  $N-1$  段階を第 2 段階と考えると第  $N-1$  段階における学習者のすべて可能な状態を (25) 式から計算する。

**Step 3.** 第  $N-1$  段階の最適決定は (27) 式で計算された **Table 1** の  $W(n, s_n)$  を最大にする決定を  $D$  から選択する。

**Step 4.** 第  $N-2$ , 第  $N-3, \dots, 2, 1$  段階を Step 1, 2 と同様の方法で学習者の状態を計算する。

**Step 5.** 全体の最適政策は (27) 式の  $W(n, s_n)$  を最大にする決定を  $D$  より選んでいく。

以上の Step に対する計算概要を示したのが **Fig. 2** のフローチャートである。

**2.3.1 実験**

**教材と装置** 単語教授システムの最適化を実際の実験より検討していくために、二種類の教授システム、 $K, J$  を設計する。システム  $K$  は最適教授プロセスを考慮した項目の提示順序をもつ。システム  $J$  は通常の直線型のブランテングで提示順序が決められる。その他、システム  $K$  では教授項目  $K$  個が 2 分割されて  $S$  個の項目の最適化に対して、システム  $J$  では教授項目  $K$  個は分割されずに直線に提示されるという違いがある。すなわち、教材は英単語 40 個の意味を教授するもので、システム  $K$  では 20 個の項目 1 と項目 2 より教

授項目がつけられているのに対して、システム  $J$  は 40 個の項目で教授する。システム  $K$  の初期のエラー確率を項目 1, 項目 2 に対して、それぞれ **Table 1** のように設定する。その結果より求めた  $W(n, s_n), d(n, s_n)$  は **Table 1** のようになり、最適政策は **Table 2** のような、提示順序になる。システム  $J$  では 40 個の項目の最適提示順序を持っていない。

システム  $K, J$  は FACOM U-200 (8k バイト) を利用してシステム化され、フィードバックメッセージを持たないプログラミングで教授項目が提示される。システムはアセンブラーを用いて記述されている。ただし、最適化はオフラインで計算され、その値にしたがってカード表示による教授項目を選定する。

実験には 6 名の東京理科大学理工工学部の学生が被験者として参加した。これらの被験者のうち 3 名は、

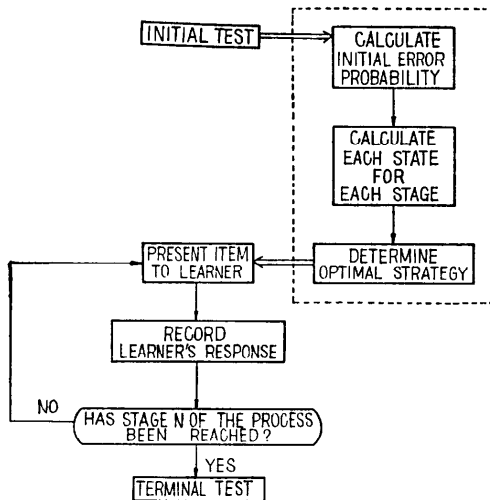
**Table 1** The iterative computation of example 2.

Number of stage	$W(n, s_n) \quad d(n, s_n)$				
	Initial State	Initial Decision	Next State	Final return	Optimal decision
1	(.03, .90)			1.06	
	(.07, .45)			1.47	
	(.15, .22)			1.62	
	(.30, .11)			1.58	
	(.60, .05)			1.34	
2	(.07, .90)	1	(.03, .90)	1.06	2
		2	(.07, .45)	1.47	
	(.15, .45)	1	(.07, .45)	1.47	2
		2	(.15, .22)	1.62	
	(.30, .22)	1	(.15, .22)	1.62	1
		2	(.30, .11)	1.58	
	(.60, .11)	1	(.30, .11)	1.58	1
		2	(.60, .05)	1.34	
3	(.15, .90)	1	(.07, .90)	1.47	2
		2	(.15, .45)	1.62	
	(.30, .45)	1	(.15, .45)	1.62	1 or 2
		2	(.30, .22)	1.62	
	(.60, .22)	1	(.30, .22)	1.62	1
	2	(.60, .11)	1.58		
4	(.30, .90)	1	(.15, .90)	1.62	1 or 2
		2	(.30, .45)	1.62	
	(.60, .45)	1	(.30, .45)	1.62	1 or 2
		2	(.60, .22)	1.62	
5	(.60, .90)	1	(.30, .90)	1.62	1 or 2
		2	(.60, .45)	1.62	

$s_1^{(1)}=0.6 \quad s_1^{(2)}=0.9 \quad \alpha=0.50$

**Table 2** The optimal sequence of item presentation

Stage 1	Stage 2	Stage 3	Stage 4
Item 1	Item 2	Item 2	Item 1
Item 1	Item 2	Item 1	Item 2
Item 2	Item 1	Item 1	Item 2
Item 2	Item 2	Item 1	Item 1



**Fig. 2** Flow diagram of list presentation for the  $N$ -stage process of instructional system

イ) システム *K, J* による学習, ロ) システム *K* による学習, ハ) システム *J* によるタスクに割当てられた。これをグループ(I)とすると、このグループによるシステムの有効性についての比較を確認する意味で、別の3名(グループ(II)とする)はタスクイ)を割当てられた。

システム*K*における最適政策は Table 1 より求めた、Table 2 のうち、第1日目→項目1, 第2日目→項目2, 第3日目→項目2, 第4日目→項目1を用いる。

被験者は各項目のなかの単語とその意味の一対ずつを順々に声を出して読む。項目の提示時間は2分間とする。その提示後に、項目のなかの単語だけを提示し、その単語の意味を FACOM U-200 のコンソールタイプより入力する。実験の最終日に最終テストを実施し、学習効果を調べる。また、その2日後に、最終テストを実施し学習効果を調べる。

2.3.2 結果および考察

**結果** グループ(I)の実験結果を Fig. 3 に示す。ただし、イ)による学習データを整理した。Fig. 3 において黒丸(●-●)はシステム*K*で、5段階で2つの教授項目を提示したときの被験者の各段階の学習状態である。この推移プロセスは最適政策にしたがって各項目を学習者に提示した結果である。Fig. 3 には最終テストの結果も点線で結んで学習効果を調べている。白丸(○-○)は最適政策による理想的な学習の推移プロセスを示している。システム*J*の結果は(◎-◎)

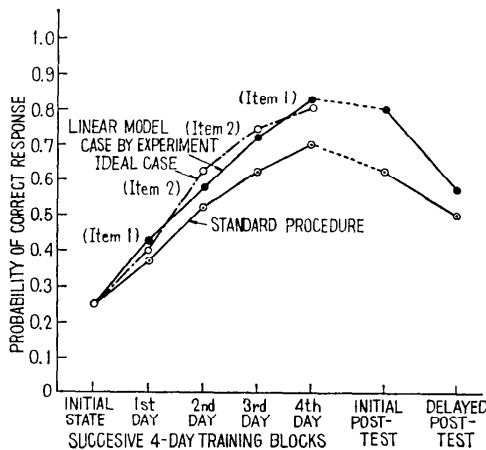


Fig. 3 Probability of correct response in linear model and standard procedure. The optimal policy in linear model is (item 1, item 2, item 2, item 1).  $s_1^{(1)}=0.6, s_1^{(2)}=0.9$

Table 3 Probability of Correct Response in Linear Model and Standard Procedure which are obtained from the results of experiment under System *K* and System *J*.

Subjects	Initial State	1st day	2nd day	3rd day	4th day	Initial post-test	Delayed post-test
No. 1	a	0.20	0.30	0.50	0.55	0.65	0.55
	b	0.20	0.25	0.32	0.40	0.45	0.35
No. 2	a	0.20	0.40	0.65	0.70	0.80	0.75
	b	0.10	0.30	0.50	0.70	0.75	0.65
No. 3	a	0.20	0.40	0.45	0.50	0.65	0.60
	b	0.20	0.30	0.40	0.50	0.55	0.45

で示めされている。グループ(II)の実験結果は Table 3 に整理されている。Table 3 の(a), (b)は、それぞれシステム *K, J* による各段階での正答率を示している。

**考察** グループ(I)のシステム*K*における結果では、最適政策にしたがったときの被験者の推移プロセスと、最適政策による推移プロセスとを比較すると、実際の状態は理想的な学習状態と類似しており、正答率は安定した増加傾向を示している。また、システム*K*と*J*とを比較すると、明らかに、システム*K*による学習状態の方が、正答率および最終テストともすぐれており、最適化政策の効果が現われていると考えられる。このことをさらに、確認する結果を示したのが Table 3 である。すなわち、いずれの被験者も、システム*K*による学習状態の方が、システム*J*よりも正答率および事後テストの成績が高いことを示している。

以上から、単語の教授プロセスのように単一オペレータモデルが適用できるときは、システム*K*による教授最適化が有効である。また、最終テストにおいて、直後の、および2日後の事後のテストの結果は、それぞれ短期記憶、および長期記憶の効果と考えられる。これらの両記憶の効果においても、単一オペレータモデルにしたがったシステム*K*の方がすぐれている。もちろん、現実の状況に適用するためには、さらに多くの被験者による検討が必要であるが、最初のグループ(I)の結果に対して、グループ(II)同様の結果が得られたことは、多段階の教授政策の可能性を明確したと考えられる。

以上で取り扱った例題は、非常に限られた教授プロセスを対象としているが、CAI システムの設計(溝口<sup>11)</sup>, 1969, 溝口ほか<sup>9)</sup>, 1973)においても、同様の考え方が暗示的(implicit)に使われている。すなわ

ち、教授論理におけるブランテングを複雑なパターンを用いたとしても、提示のループはブランテング内のカウンター数に依存している。したがって、カウンターのループ数だけ、少なくとも学習者はエラー確率を減らしていくことができる。この視点から考えると、学習者の認知状態などを意識しなくてもよいトレーナー、単語学習の CAI では、ブランテングカウンターの最適化が教授最適化のひとつの方向であることを示している。

2.4 学習者のペイオフ問題

2.4.1 問題の設定とシミュレーション

学習者が教授プロセスとのかかわりの中で、学習者の自信の度合から生ずる期待値を考慮しなければならない状況がある。たとえば、学習者への教授項目についてはいくつかのレベルがあり、それに対する評価点も、レベルごとに変るときに、学習者はどの教授項目を選択して、学習していくかという場合である。おそらく、レベルの高い教授項目を学習していけば、学習者は高い達成感とスコアとを得て、学習に対する意欲も満足感ともに増大していく。逆に、その選択を誤まれば自信を失ない、達成感も損なわれてしまう。こうした学習者の行動には、学習プロセスの原理というよりもむしろ学習者の期待値原理によって説明される場合が多い。すなわち、学習者がある程度まで自分自身を評価して、その見積り、すなわち期待値に対応する教授項目を期待値の最大化として選択するペイオフ問題に帰着される。このことは、教授システムを構成する側についても当然、考慮する必要のある問題である。例えば、この教授項目のレベルには、だいたい学習者の 60% が選択するので、スコアとしてはだいたいこのくらいの値を見積っておけばよいだろうという場合である。しかし、ここでは、前者の学習者側の問題を設定し別の教授最適化の視点を明確にしていく。

問題 学習者は 4 日間で、2つの教授政策のうちのいずれかを選び学習する。政策 I では、教授レベルが高く、したがってテストもむずかしいが、スコアの配分は政策 II と比較して、正答、誤答いずれの場合も 2 倍になっている。したがって、正答のときは高いスコアが得られ、誤答のときは低いスコアとなってしまう。政策 II では、政策 I とは逆で、教授レベルは高く設定されていない。また、スコアの配分も政策 I の 1/2 である。各段階のテストスコアは政策 II を基準として 10 点とする。ただし、政策 I を選択したときには、選択した意欲に対して 5 点が加算されるので、正

答のときは 25 点を獲得できる。4 日後に、事後テストが実施され、正答のときは 30 点が加えられる。

学習者は次の 2 つのうちのいずれかの状態にある。

$$\begin{cases} S=1 & \text{教授項目へのテストに正答} \\ S=2 & \text{教授項目へのテストに誤答} \end{cases}$$

そして、教授政策 (I, II) へのオプションを次のようにする。d=1 政策 I で学習、d=2 政策 II で学習する。その他、特に重要なのは、状態推移の確率で、Table 4 で設定した値を用いている。基本的な考え方は、政策 I では高レベルの教授内容なので、たとえば、s<sub>n</sub>=1 (正答) であったとしても、次の s<sub>n-1</sub> が正答となる確率は 0.4 と低い値で、誤答となる確率は 0.6 である。逆に、政策 II では、s<sub>n</sub>=1 であれば、次も正答であるという確率は 1 であるものとする。Table 4 にはペイオフ値 g<sub>n</sub>(s<sub>n</sub>, d) も示してある。これらの値を用いて計算したのが Table 5 である。小数点以下の数値は丸めてある。Table 5 の d(n, s<sub>n</sub>) に最適政策を示している。第 1 日目は、政策 II で確実に点を獲得し、あと n 日に対しては、各プロセスの期待値を計算し、政策 I か II かを決定する。

Table 4 Structure of criterion

Instructional Decision	P(S <sub>n-1</sub>  s <sub>n</sub> , d <sub>n</sub> )	Expected value	
d <sub>n</sub> =1	s <sub>n</sub> =1	s <sub>n-1</sub> =1    s <sub>n-1</sub> =2	g <sub>n</sub> (1,1)=25
		0.4    0.6	g <sub>n</sub> (1,2)=-15
	s <sub>n</sub> =2	0.2    0.8	
d <sub>n</sub> =2	s <sub>n</sub> =1	s <sub>n-1</sub> =1    s <sub>n-1</sub> =2	g <sub>n</sub> (2,1)=10
		1.0    0.	g <sub>n</sub> (2,2)=-5
	s <sub>n</sub> =2	0.7    0.3	

Instructional decision  $D = \begin{cases} 1 & \text{for strategy I} \\ 2 & \text{for strategy II} \end{cases}$

State of Learner  $S = \begin{cases} 1 & \text{for correct response} \\ 2 & \text{for incorrect response} \end{cases}$

Final value  $W(0,1)=30$   
 $W(0,2)=0$

Table 5 Iterative computation of dynamic programming

n	s <sub>n</sub>	d <sub>n</sub>	W <sub>d<sub>n</sub></sub> (n, s <sub>n</sub> )	W(n, s <sub>n</sub> )	d(n, s <sub>n</sub> )
1	1	1	25+0.4(30)+0.6(0)=37	40	2
		2	10+30=40		
1	2	1	-15+0.2(30)+0.8(0)=-9	16	2
		2	-5+0.7(30)+0.3(0)=16		
2	1	1	25+0.4(40)+0.6(16)=50.6=51	51	1
		2	10+40=50		
2	2	1	-15+0.2(40)+0.8(16)=5.8=6	28	2
		2	-5+0.7(40)+0.3(16)=27.8=28		
3	1	1	25+0.4(51)+0.6(28)=62.2=62	62	1
		2	10+51=61		
3	2	1	-15+0.2(51)+0.8(28)=17.6=18	39	2
		2	-5+0.7(51)+0.3(28)=39.1=39		
4	1	1	25+0.4(62)+0.6(39)=72.4=72	72	1 or 2
		2	10+62=72		
4	2	1	-15+0.2(62)+0.8(39)=28.6=29	50	2
		2	-5+0.7(62)+0.3(39)=50.1=50		

### 2.4.2 学習者のペイオフ問題と教授最適化

上述の問題では具体的教授システムへの応用例までを考えていないが、学習者のもつペイオフから各段階にどのような教授政策を選択していけばよいかを考えた。その意味では、さらに、実際のシステムを構成する必要があるが多段階教授プロセスを別の視点から接近したと言えよう。通常、学習者の期待値はその学習者の主観的確率をどのように評価するかという問題になり、人間の意志決定論(宮沢<sup>10</sup>, 1971)の枠組に入ってくる。さらに、さまざまな教授コースの選択が学習者に許されたときに、どのような要素から教授コースの選択を行うか、また、その要素間の代数的結合ルールを数量化の方法(林<sup>6</sup>ら: 1973, 永岡, 藤田<sup>16</sup>ら: 溝口<sup>15</sup>: 1974)で解析することも可能である。

次に、教授最適化の概念を拡張するために、教授政策に要するコスト、とくに、学習者の理解状態を診断するに要するコストを導入する必要がある。このとき、教授最適化は多段階の各段階の期待値を最適化する問題から、学習者の理解状態を診断するコストを最小化する問題に展開することができる。すなわち、学習者の理解状態が  $k$  段階にあるときに、 $i$  段階にあるとしたときの診断ミスのコスト  $C_{ik}$ 、学習者のさまざまな理解状態  $h_i$  の分布  $H$  への診断コストを  $C(H)$  とする。状態  $h_i$  にある確率  $P(h_i)$  とすると、 $C(H)$  は次式のように記述でき、診断コストを導入したときの定式化ができる。

$$C(H) = \min_k \sum_i P(h_i) \cdot C_{ik} \quad (28)$$

実際には、学習者の応答  $r_j$  とすれば(28)式は次のように書ける。

$$C(H|r_j) = \min_k \sum_i P(h_i|r_j) \cdot C_{ik} \quad (29)$$

したがって、教授プロセスで何らかの質問  $Q$  の結果の後に、 $r_j$  を得たときの期待コストは次式で表わされる。

$$E[C(H|Q)] = \sum_j \left\{ P(r_j) \min_k \left[ \sum_i P(h_i|r_j) C_{ik} \right] \right\} \quad (30)$$

こうした定式化より、質問  $Q$  への学習者の応答  $r_j$  を用いて、 $i$  段階の理解状態にあるときに  $k$  段階とミス診断するコストを最小化することに接近することが可能である。ここでは診断の論理で使われるベイズの定理の導入(溝口, 佐伯<sup>17</sup>: 1972, 佐伯, 溝口<sup>18</sup>: 1973)をとくに考えていないが、CAIの予備フレームから学習者に適合するフレームを診断する方法に kullback の理論を導入する竹谷, 佐藤<sup>22</sup> (1974) の研

究方向も考えられる。

### 3. むすび

本論文は多段階の教授プロセスの最適化を動的計画法にもとづいて定式化した。動的計画法は、次元が高くなると計算量が膨大になるという“次元の呪い”を持つ。教授最適化の対象を線形モデルにもとづくエラー確率として動的計画法を適用すると、現実的に単語教授システムに利用できることを例題と実験で示した。次に、教授評価が期待値の形式で表わされる時の教授最適化を、学習者のペイオフから検討し、例題のシミュレーション解析で議論した。さらに、学習者の期待値、主観確率の問題を発展して、教授診断コストの概念を導入し定式化した。こうした方向は、教授プロセスをさまざまな分野で開発された方法、例えば、数量化法、情報理論、意志決定論、さらに人工知能論(carbonell, J. R., 1973<sup>3</sup>)などを積極的に取り入れていくことによって教授最適化をめざすものである。

終りに、本論文は著者が東京大学工学部電子工学科宇都宮敏男教授および同大学医学部、医用電子施設斎藤正男教授のもと内地留学した際に書かれたものである。日頃、御鞭達頂いた宇都宮、斎藤両教授に感謝するとともに、内地留学を許可していただいた、東京理科大学理工学部経営工学科主任上甲教授をはじめとする諸先生方に感謝する次第である。また、本研究は昭和47年度松永研究助成金が発端になっており、同財団に感謝する。

### 参考文献

- 1) Atkinson, R.C. & Paulson, J.A.: An Approach to the Psychology of Instruction, Psych. Bull., (1972), 78, pp. 40~61.
- 2) Bruner, J.S.: Some Theorems on Instruction State with Reference to Mathematics. In E. R. Hilgard (Ed.), Theories of Learning and Instruction: The Sixty Third Yearbook of National Society for the Study of Education, Chicago; The University of Chicago Press, pp. 306~335 (1964)
- 3) Carbonell, J.R. & Collins, A.M.: Natural Semantics in Artificial Intelligence, Proceeding of IJCAI3, pp. 344~351, Stanford, California (1973)
- 4) 藤田広一, 小沢慎治ほか: CAI タイプトレーナーについて, 電子通信学会教育技術研究会資料, 資料番号 E 72-4, pp. 89~97 (1972)
- 5) Groen, G.J. & Atkinson, R.C.: Models for Optimizing the Learning Process, Psychological



- Bulletin, Vol. 66, pp. 309~320 (1966)
- 6) 林, 樋口, 駒沢: 情報処理と統計数理, 産業図書 (1973)
  - 7) Kopstein, F. F. & Seidel, R. J.,: Computer as Adaptive Instructional Decision Maker, Professional paper 1-74, HumRRO, Division No. 1, Alexandria, Virginia (1970)
  - 8) Matheson, J.,: Optimum Teaching Procedures Drived from Mathematical Learning Models, EES Tech. Report No. CCS-2, Stanford Univ. (1964)
  - 9) 三重野博司, 溝口文雄: コンピュータシステム論, オーム社 (1973)
  - 10) 宮沢光一: 情報, 決定理論序説, 岩波 (1971)
  - 11) 溝口文雄: 教育自動化における人間-機械インターフェイス, 人間工学, Vol. 5, No. 6, pp. 389~396 (1969)
  - 12) 溝口文雄: システムの機能, 坂元昂編講座教育のシステム化I, 教授学習過程のシステム化, pp. 23~45, 明治図書 (1971)
  - 13) 溝口文雄, 佐伯胖: CAI 教授論理と決定機構, 電子通信学会教育技術研究会資料, E 72-4, pp. 71~80 (1972)
  - 14) 溝口文雄, 佐伯胖: CAI 教授論理と学習者意志決定機構, 情報処理, Vol. 15, No. 2, pp. 101~109 (1974)
  - 15) 溝口文雄: Conjoint Measurement 解析による多次元尺度構成へのアプローチ, IE review, Vol. 15, No. 3, pp. 117~123 (1974)
  - 16) 永岡慶三, 藤田広一ほか: SP 表 (スケエログラム) の教育評価への応用, 第1回行動計量大会論文抄録集, pp. 116~117 (1973)
  - 17) Neisser, U.,: Cognitive Psychology, Appleton-Century-Crofts; New York (1967)
  - 18) 佐伯胖, 溝口文雄: ベイズ診断ロジックによるCAI システム, 電子通信学会教育技術研究会資料, ET 73-7 (1973)
  - 19) 佐藤隆博: CAI について, 情報処理, Vol. 11, No. 3, pp. 155~167 (1970)
  - 20) 佐藤隆博: CMI について, 情報処理, Vol. 12, No. 5, pp. 368~379 (1973)
  - 21) Smallwood, R. D.; The Analysis of Economic Teaching Strategies for a Simple Learning Model, Journal of Mathematical Psychology, Vol. 8, pp. 285~301 (1971)
  - 22) Takeya, M. & Sato, T.,: Construction and Evaluation Methods for Instructional Programs in CAI, NEC Research & Development, No. pp. 31~38 (1974)

(昭和50年9月18日受付)

(昭和51年10月4日再受付)