制御行動における適応と学習

黒川 隆夫*
田村 博

1. はじめに

制御とは簡単にいえば、対象とする系が望ましい動作をするように、この系を駆動することである。系を駆動する信号すなわち操作量を人が与える場合、この制御系は手動制御系とよばれる。また手動制御系において人の果たす役割を機械すなわち制御器に置き換えたものが自動制御系である。

制御系は制御対象とそれを駆動する制御器（または人間オペレータ）より成り、通常図1のようなブロック線図で表わすことができる。ここで r は制御対象の望ましい動作を示す目標値、また x は制御対象の実際の動作である制御量（応答）を示しており、両者の偏差 e の関数として操作量 u が決定される。ここでは、制御は偏差 e を 0 に保つように制御行動を実施する（操作量を発生する）ことである。

手動制御系においては従来人が果たしている役割を理解する目的で人間オペレータの伝達関数が測定されてきた。これらの結果は人の制御動作の平均的な性質を与え、人間—機械系を設計するある程度の目安を与えてきた。また目標値や制御量の変化がかなり速くて、人間の動作特性が物理的あるいは生理的な因子によって制限されている場合には伝達関数における無駄時間と時定数というパラメータで人間の動作特性をある程度記述することも決して無理ではないと思われる。

しかし多くの場合、手動制御系の応答と人間を等価な伝達関数で置き換えた自動制御系の応答はかなり異なっている。そしてこの相違は人間オペレータのランダム性あるいはノイズによるものであるという形で処理されてきた。

両制御系の違いの1つは制御対象が複雑になるに従って、人が非線形的な on-off 動作 (bang-bang 動作

ともよばれる) あるいはパルスのな動作によって目的を達成しようとする傾向をもつ点である (図2参照)。また人が制御状況の変化に対して適応性と学習機能をもっていることも相違点としてあげられる。

図1. 制御系のブロック図

図2. 手動制御にみる bang-bang 動作。(c) は操作量の変化を表わすが、中間値をとれるボリュームを用いているにもかかわらず両端値のみを用いる bang-bang 動作を実行している。制御対象は3次系。

制御対象の複雑さには種々のものが考えられるが、ここでは動特性を表わす微分方程式の次数が高次になるほど複雑であると考え。この次数が3になると、普通はある程度の練習を積み重ね限り制御することはできないし、4次になるとまず不可能であるといつてよい。このような次数に関する制御の限界は決して人間の筋肉の応答速度や感覚器の識別限界によって与えられるものではない。

制御対象の複雑化とともに人間の制御動作のパターンが変化することについての1つの解釈は、動特性の次数が高くなるに従って制御動作の結果が制御対象の応答に直接反映されなくなり、制御対象の性質を認識したり操作の適否を判断することが困難となるためという考え方である。そして操作量を図2(c)のように bang-bang 型に選ぶことによってこの不利を補っているとみなせる。さらにこれとは別に、bang-bang 制御が制御に要する時間 (制御時間) を最短にしたり、

Adaptation and learning in control behavior:

* Takao KUROKAWA, Hiroshi TAMURA 大阪大学基礎工学部生物工学科 (Department of Biophysical Engineering, Faculty of Engineering Science, Osaka University)

平均偏差を最小にするうえでもっとも都合のよい制御方式であることが理論的に知られており、人間オペレータは無意識のうちにこのような制御方式を実行しているのであるとみることもできる。

ともかく、ある対象を制御できるためには、制御対象の特性についての理解を深めることが不可欠である。このために人間は自分の行なった制御動作のパターンとそれの直接の反映として制御対象の動作のパターンを認識するとともに、両者を総合することによって対象の性質を認識しているのであろう。この認識が進むとともにこんどは制御動作を改善することが可能となる。前述の bang-bang 型の制御は制御動作の改善の結果、人間オペレータが習得する制御方式である。

たんに手動制御だけでなく、広く生物の行動は一般にパターン認識とパターン発生に基づいており、これから生体の複雑な動作様式が可能になっていると考えられる。たとえば、われわれが日常何気なく実行している手足の動作はこのようなパターン処理に基づいた不断の制御過程の結果であるといえる。現在のロボットの手足の運動にはパターン認識とよべるものがほとんど取り入れられていないが、ロボットが自由度の多い手足の運動を生き物のように行なえるようになるにはパターン認識制御といったものが必要と思われる。

本文ではこのようなパターン認識制御の可能性について述べるわけであるが、その前に従来の適応制御の方式を簡単に振り返ってみることにする。

2. 適 応 制 御

生物が適応性をもつことは衆知であるが、適応の概念は工学にも導入され、たとえば「不十分な事前情報を行動過程で補ないながら系が自己の構造を変化させて、系に課せられたある目的を達成することである」と定義されている。このような適応機能を必要とする工学系としては制御系が代表的である。図1に示した制御系に制御のよさに関する評価法が導入された場合、最良な制御は最適制御とよばれ、技術の発展とともにその地位は高まりつつある。目標値、制御対象の動特性および制御の最適性の評価法が完全に与えられる場合には、決して容易ではないが何らかの方法により最適制御法を見つけることが可能である。しかしながら現実の制御対象は動特性が厳密には知られていないのが普通であり、このとき解析的に最適制御問題を解くことは不可能となる。

ここに適応性の導入が望まれるわけで、与えられた

情報の不確定さを補うために制御過程で情報を収集して系の構造を改善し、最適性を追求する種々の適応制御系や学習制御系が提案されている。図3の2つのブロック線図は従来の適応制御系の代表的な型を示したものである。

図 3. 従来の適応制御系

図3 (a) は外部モデル型とでもよぶべきものであり、最適制御を実行する原型制御器（ときには人間がこの役割を果たす）があって、その動作を見習って適応制御器が最適な制御動作を学習してゆく。この原型制御器は最適制御動作のモデルであって、適応制御器は一種の物まね機械と考えられる。この適応制御器は学習期間中は制御対象を含む制御ループとはオフラインの関係にあり、学習完了後にモデルに代って制御行動を実行する。未知の制御対象に対してはこのような原型制御器の存在が問題となる。

図3 (b) は (a) のように既成のモデルを利用するのではなく、適応制御器が自力で最適制御動作を見つけることができるより一般的な適応制御系である。前述のように制御対象の動特性が明らかになれば最適制御動作を見つけることができるので、この系では制御対象の数学的モデルを仮定し、そのパラメータを制御過程で得られる情報から同定することに重点を置いている。

これらに代表される現在の制御技術においては、各信号路を流れる情報が数量として捕えられ、それらについての数学的な操作結果に基づいて制御系が動かされている。

一方人間の手動制御あるいは広く生物の適応行動においては、工学的に用いられているような数学的モデルや制御の最適性に関する評価法に基づいた適応制御方式が用いられているとは考えられない。人間オペレータは制御動作の結果得られる応答のなかから制御に有効な情報を定性的なパターンとして抽出することを制御の初期に自力で学びとり、その後自分の制御動作に対する一定の応答パターンを予測し、これと実際の応答パターンとの比較を通じて制御対象についての認識を深めると同時に目的を達成するのに必要な制御動

作を見出してゆくものと考えられる。そして生物のすぐれた適応と制御の能力が生体の得意とするパターン認識の機能に本質的に基づいていると考えるのは自然であろう。

パターン認識を制御と結合させることによって、より簡単な論理に基づく高度な適応と制御の機能が工学的にも実現されるかどうかは非常に興味ある問題であり、以下ではその可能性の1つを簡単に示したいと思う。これはまた1.で述べた手動制御系における人間オペレータのモデルと考えることもできる。

3. 最短時間制御

最適制御の重要なクラスの1つとして最短時間制御がある。ここではこの最短時間制御を2.で述べたような従来の手法とは異なり、生物が行なっていると考えられるようなパターン識別に基づいた論理に従って適応的に実現することを考える。

最短時間制御は制御対象の制御量 $x=(x_1, \dots, x_l; \dot{x}_i = x_{i+1})^*$ を可能な最小の時間で目標値 $r=(r_1, \dots, r_l; \dot{r}_i = r_{i+1})$ に移すことが目的であり、操作量 u が $|u| \leq 1$ のように制限されている場合には、 $u = \pm 1$ の2値を適当な論理に従って切り換えることにより実現できることが知られている¹⁾。操作量の変化として表わされる制御動作を時間関数 $u(t)$ として与える方法と偏差の関数 $u(e)$ として与える方法があるが、安定性の点からは図1のような閉ループ制御系として構成される後者の方式がすぐれている²⁾。このとき偏差 e の関数として制御動作を行なう制御器の機構は偏差空間 $E=\{e\}$ における操作量の切り換え面で表わされる。図4は3次の制御対象（したがって偏差空間は3次元）に対する最適切り換え面とこれによって制御される偏差 e の遷移軌道の関係を模式的に描いたものである。制御量を目標値に移すことは、任意の偏差 e を o にすることすなわち e を空間 E の原点に移すことに相当する。この空間 E において、最適すなわち最短時間制御を可能とする切り換え面はいくつかの特徴的性質をもつ³⁾。いまだか $i-2$ 回の操作量の切り換えで原点へ最適制御できる e の集合を第 i 次の切り換え面と名付けると、各次の切り換え面は一般に2枚の滑らかな超面の結合として表わされ、それより高次の切り換え面に含まれるという入れ子構造を成して

いる。そして任意の偏差 e は各次の切り換え面によって操作量が切り換えられてつきつき低次の切り換え面上に移され、最終的には第1次の切り換え面である原点に達する。

図4. 最適切り換え面の模式図（3次系の場合）

以上は最適な切り換え面についての話であるが、制御対象の動特性が未知であると最適切り換え面は解析的に求められない。この場合に最適制御を達成するには空間 E に切り換え面を任意設定してこれを用いて実際に制御を行ない、この間に得られる情報によって切り換え面の構造を逐次修正して最適切り換え面に近づけることが考えられる。この適応制御のための論理は切り換え面に対する偏差の相対的な遷移パターンを解析することによって導かれる。図5は2次元偏差空間における最適およびそうでない切り換え面を描き、それぞれの場合にみられる偏差 e の典型的な遷移パターンを示したものである。(a)は最適切り換え面の場合であり、偏差が面上に達すると操作量の切り換えが生じて切り換え面上を目標である原点に向かって進む。これに対し、切り換え面が最適でない場合には偏差が切り換え面から離れたり(b)、切り換え面上を操作

図5. 偏差の遷移パターンと切り換え面の修正

* $\dot{x}_i$ は x_i の時間微分を表わす。 l は動特性の次数。ゴダック体はすべてベクトル量を表わす。

量の頻繁な切り換えを伴いながら滑る (c) のが観察される。図 5 (b), (c) の矢印はこのような遷移パターンが生じたときに切り換え面を修正すべき方向を示しており、この修正を繰り返せば (a) にみられるような最適制御動作を与える切り換え面に収束することが確認されている。⁴⁾ 制御対象が高次系になってもまったく同様である。

切り換え面の構造修正は数学的には学習機械にみられるのと類似のアルゴリズム^{5), 6)}に従ってなされるが、適応制御器自身が修正方向決定機構を備えていて外部からは何らの教示も必要としないため、このアルゴリズムによる学習を教師なし学習または自己学習とよぶことができる。⁶⁾

4. 自己適応制御過程

偏差空間における偏差の遷移パターンを情報とする以上の自己学習アルゴリズムを備えた適応制御系は図 6 のブロック線図によって表わされる。ここで単線の信号路はスカラー量を、複線はベクトル量を示している。

図 6. パターン識別法を用いた自己適応制御系

さらに制御器を貫く矢印は制御器すなわち切り換え面の構造に変化を与える学習の効果を表わしている。

図 4 でみた最適切り換え面の入れ子構造への精度よい収束を得るために多項式から成る階層構造をもつ切り換え関数によって切り換え面を表現する。^{3), 4)} 空間 E に 1 つの初期偏差が与えられてのち、原点へ制御するまでを制御の 1 ステップと数え、前節の自己学習アルゴリズムを導入したときの適応制御過程を図 7 に描く。図中に示した 2 次の制御対象を未知と仮定したもので、図 7 (a) には切り換え面の変化 ($e_2 \geq 0$ のみ) を、(b) には学習曲線に相当する制御成績が改善される様子を描く。ただし第 m ステップでの制御成績 q_m は

$$q_m = \frac{100}{m-k+1} \sum_{j=k}^m \frac{t_j - t_{0j}}{t_{0j}} (\%),$$

$$k = \begin{cases} 1 & m \leq 14 \text{ のとき,} \\ m-14 & m \geq 15 \text{ のとき} \end{cases}$$

で定義する。ここで t_j は第 j ステップでの制御時間、

t_{0j} は同じ初期偏差に対する最短制御時間である。したがって連続した 15 ステップにおいて最短時間制御が実現されると $q_m = 0$ となり、 q_m の値が 0 に近いほど最適に近い制御が行なわれたことを示す。図 7 より、この適応制御器は急速に自己の制御動作を改善して最短時間制御に近い成績をあげられるようになることがわかる。

図 7. 適応制御過程の 1 例 (2 次系)

この例は未知ではあるが一定の動特性をもつ対象への適応を示したものである。一方目標値が時間変化をしたり^{*}、環境の影響などにより制御対象の動特性が変化する場合にも同様な適応過程が観察される。このとき目標値や動特性などの変動により新しく発生した状況への適応は過去の適応成果の消失を意味するので、ふたたび同じ状況が生じた場合にも改めて適応手続きを必要とする。しかし制御状況の変動が適応速度に比べて遅ければ、以上のような適応制御によって本来の目的を十分達成できる。

5. 学習制御と連想制御⁷⁾

制御状況の変化が速くなると、以上のような適応制御方式では切り換え面の適応的最適化が状況の変化にもはや追従できなくなる。温度、湿度、風速、荷重の大きさなどの物理量によって記述できる環境状態の変化が制御対象の動特性を変化させる原因となることはしばしば観察されることである。このように制御対象の動特性変化の原因が環境変化にあると仮定し、3 つの環境状態 ($z = 0, 1, 2$) がランダムに生起する場合の適応制御過程を制御成績の変化によって表わしたのが図 8 (a) である。未知対象として図中に示した 2 次系を用いたが、制御成績は一定範囲内で振動を繰り返すのみである。

クレーンの操作は 1 つの手動制御であるが、人間オペレータはクレーンの荷重の大きさによって制御動作

^{*} 同じ最適切り換え面を規定する時間関数目標値は同じものであると考える。図 7 の例では $r_1(t) = (at + \beta, \alpha)$ と $r_2(t) = (rt + \delta, r)$ は同じ目標値となる (α, β, r, δ は任意の定数)。

を変えることを要求される。この場合人間オペレータは新しい荷重量に対する制御に慣れても、以前に扱った荷重の大きさに対する制御法をある程度記憶していて、ふたたび同じ大きさの荷重を扱わねばならなくなったときにはただちに適切な制御を行なうことができる。また経験が豊富になるにつれて制御動作も上達したものとなる。このように新しい情況へ適応しても過去に遭遇した情況への適応成果を何らかの形で蓄積できて、再度同じ情況が発生したときにその成果を活用できる機能を備えた制御系は学習制御系とよばれる。前述のような情況に対処できるためにはこのような機能が不可欠となる。高等動物の動作は一般に学習制御方式に基づいているとみなせる。

図 8. 適応制御と学習制御

偏差空間 E を環境状態空間 $Z = \{z\}$ さらには目標値を特徴づけるパラメータ空間 Y を含むように拡大して新しく状態空間 $\mathcal{E} = \{e\}$ を構成し^{*)}、ここに切り換え面を設定することによって前節の適応制御系は学習制御系に拡張できる。状態 e は観測可能であるとする。切り換え面の拡張された部分は適応成果の記憶領域として動作し、情況が変化すればただちにに対応する切り換え面がよび出され、過去に達成された適応段階から適応過程を再開することができる。切り換え面の修正には3. で述べた自己学習アルゴリズムが用いられる。

この学習制御系は図9のように表わされ、図8(a)を得たのとまったく同じ条件下で未知対象を制御した結果を図8(b)に示す。ランダムに発生する3つの環境状態に対する最短時間制御が同時に実現される過程を読み取ることができ、適応制御との差が顕著に現れている。図10にはこのときの切り換え面の変化と最適切り換え面を描いた。

^{*)} 状態空間 $\mathcal{E}$ は E と Y と Z の直積空間となる。

図 9. 学習制御系

図 10. 学習制御による切り換え面の変化

生物のもつ認識機能に連想がある。これは2つの概念間の距離を何らかの尺度で測って近いものを相互に関連させあうことであり、この能力のおかげで生物は随分得をしている。新しい情況が生じたとき、過去に経験したものからこの情況に類似のものを選び出して、それに対応する最適動作を実行することにより新しい情況への適応を一から始める無駄を省くことができる。この学習制御器にもこれと類似の機能が見出せる。図10より明らかに、経験環境だけでなく未経験環境に対する切り換え面の最適化が行なわれている。この結果未経験環境での潜在的な制御能力も向上し、新しく発生した環境状態に近接する環境状態に対してすでに最適な制御動作が確立されているならば、新環境においてもほぼ最適な制御を実行できる。このような機能は連想制御という名でよぶことができよう。

6. 制御系の超安定性

手動制御系において制御対象の動特性が突変すると、人間オペレータはまず制御系を安定にすることに努力を払い、しかるのちふたたび制御動作を改善して最適制御を達成するようになることが観察されている。制

御系では何にもまして系の安定性が確保されていなければならない、制御動作の最適化はそのあとで達成されるべき二義的な機能といえよう。この意味で人間オペレータの行動様式はきわめて適切なものである。述べてきた適応制御器・学習制御器も人間オペレータと同様に安定性の面においてもすぐれた性質を示す。

図11は切り換え面が最初、系の動作を不安定にする位置にあるときに、前述の自己学習アルゴリズムを適用した場合の適応制御過程を描いたものである。すなわち図11 (a) の AO を切り換え面とする場合には、ステップ目標値に対して偏差成分 e_1 は図11 (b) の1点鎖線で示されるように振動的に発散して $e=0$ への制御は不可能となる。ところが自己学習アルゴリズムを導入すると、図11 (b) に実線で示すように、はじめ約60秒間は発散的であるが、その後次第に原点に収束する。これは切り換え面が適応修正を受けて BC のように変化し、制御系に安定性がもたらされたためである。

図 11. 自己学習アルゴリズムによる制御系の安定化

制御の第2ステップにおいても同じ初期偏差が与えられたとすると、図11 (b) の破線のような応答が得られ、制御時間は第1ステップでの約28分の1に短縮されている。系の振動性はほとんど抑えられ、安定化が最初のステップでほぼ達成されたことがわかる。

W. R. Ashby は、生物系は種々のフィードバック制御機構によってホメオスタシスを維持しているが、環境変動が大きくてこの機構のみではホメオスタシスの維持が困難となった場合に、能動的な適応探索を実

施して新しい環境にあったホメオスタシスを確立してゆく能力をもつことを指摘し、この安定性を獲得する性質を¹⁰⁾超安定性とよんだ。上述の制御系の機能はちょうどこの超安定性に相当するものである。

このように1つの最適性を追求する適応機構が、安定化機能を同時にかね備えていることは制御系にとってきわめて好都合なことであり、3. に述べたようなパターン識別法を用いることによって実現された注目すべき性質である。図11の例によっても示されたように、安定化の過程は適応過程の初期に存在し、制御系における必須条件である安定性をまず確保するという制御の基本原理にもなっている。

7. おわりに

パターン認識の結果に従って制御を実行するとともに制御動作の改善も行なうパターン認識制御の1つの可能性を、主として手動制御系における人間オペレータの動作と対比させながら述べてきた。従来なされてきたように人間オペレータを伝達関数という堅い機械に等価なものとして捕えるのではなく、適応・学習・判断のできるパターン認識系として把握することが人間の行動を理解するうえで必要である。ここで示したパターン識別法を用いる適応制御と学習制御は初歩的なものであるにもかかわらず、種々のすぐれた機能を備えており、パターン認識制御の有用性が示唆されている。

生物の行動については触れることができなかったが、1. で述べたように生物の行動はパターン認識およびパターン発生とは不可分であり、しかもこれらが制御という機構を通して結びついていると考えられる。本能や無条件反射は一定の刺激パターンが引き金の役目を果たす1種のプログラム制御であり、意識的に行なわれる日常行動の多くはさらに複雑なパターン認識に基づいた多変数フィードバック制御の結果である。さらに、技術的には主として静的パターンの認識が議論されているが、生物が動的なパターンの定性的な認識にすぐれた能力をもっていることは注目すべき点である。

「制御は生命の起原と同時に始まった」というV. A. Trapeznikov¹¹⁾の言葉は決して誇張ではなく、人間をはじめとする生き物から学ぶべき制御技術はまだ多いと思われる。

文 献

- 1) L. S. Pontryagin, W. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mishchenko: "The Mathematical Theory of Optimal Processes" John Wiley, New York (1962); M. Athans, P. L. Falb: "Optimal Control" McGraw-Hill, New York (1966)
- 2) 高橋安人: "システムと制御" 岩波書店, 東京 (1968)
- 3) 黒川隆夫, 田村博: 「電子通信学会論文誌」C 53-C, 904 (1970)
- 4) H. Tamura, T. Kurokawa: Adaptive Pattern Classifiers Applied to Time Optimal Control, "Technical and Biological Problems in Control" Instrument Society of America, Pittsburgh (1970); 岩本哲夫, 黒川隆夫, 田村博: 「第13回自動制御連合講演会」127 (名古屋, 1970年10月)
- 5) N. J. Nilsson: "Learning Machines" McGraw-Hill, New York (1965)
- 6) Я. З. Пыцкий: "Адаптация и Обучение в Автоматических Системах" Издательство «Наука», Москва (1968)
- 7) 黒川隆夫, 田村博: 「電子通信学会論文誌」C (予定).
- 8) J. E. Gibson: "Nonlinear Automatic Control" McGraw-Hill, New York (1963)
- 9) D. C. Miller, J. I. Elkind: *IEEE Trans. Human Factors Electron.*, HFE-8, 218 (1967); A. V. Phatak, G. A. Bekey: *IEEE Trans. Man-Machine Sys.*, MMS-10, 72 (1969)
- 10) W. R. Ashby: "Design for a Brain" John Wiley, New York (1960)
- 11) V.A. Trapeznikov: 「制御工学」7, 2 (1963)

■ 第11回生化学若い研究者の会夏の学校案内 ■

大学院生, 助手を中心とする若手研究者の皆さん, 第11回「生化学若い研究者の会夏の学校」で, その若き, そして燃ゆる情熱をぶっつけあい, 私達若い世代で生化学の明日を切り拓いていく出発点にしようではありませんか。

窒息するような研究室の窓を大きく開け放って, 乗鞍高原の清新な大気と全国の仲間の熱い息吹を流れこませようではありませんか。全国の若い研究者達と, 生化学の未来の問題, 自分の携っている研究の問題, 研究室の現状等々を, 夜を徹しても語り合い, 交流を深め, 親睦を図ろうではありませんか。

「生化学若い研究者の会」は, 今から11年前に「良い研究をしたい」「全国の仲間と研究について語りあいたい」という若手研究者の声があつまって誕生しました。以来「夏の学校」では, 個人個人の研究を紹介する小分科会や生化学の今日的な中心問題を深めるシンポジウムをはじめ, 研究の内容を知り深める活動, そして研究を生き生きとやっていくために, 研究室をもっと研究しやすい場にしていこうとする研究体制改善の活動を車の両輪として活動してきました。

全国の若い研究者の皆さん, さあ共に乗鞍高原に集まろう。

会 期 1971年8月7日(土)～8月11日(水)

場 所 乗鞍高原(長野県南安曇郡安曇村)

内 容

- ・総会
- ・シンポジウム(テーマ・生体とその環境変化)
- ・研究体制シンポジウム

[以上, 全体討論会]

- ・分科会(蛋白質の構造と機能, 細胞と細胞の相互作用, 細胞の表層構造と細胞の特異性, 核酸の高次構造, 酵素の localization, 細胞内顆粒の生合成と分解, 膜の脂質生合成, 膜透過, 進化のみちすじ, 生化学における技術, 実験材料・実験系, 高等植物の生化学, 分子遺伝学の問題点, 発生と分化, その他若干のテーマ)

[以上, 30～40人位の分散討論会]

- ・小分科会(夕食後10～13人単位で自分の研究紹介および討論や研究上の問題点等々について話しあう.)
- ・問題別集会(現在の日本の研究体制のひずみや公害等々の問題テーマについて討論する.)
- ・自由集会

「夏の学校」の詳細について知りたい方, ご要望がありましたら下記までご連絡ください。

京都市左京区北白川追分町
京都大学理学部化学教室
生 物 化 学 研 究 室 内

生化学若い研究者の会夏の学校事務局
TEL 京都 (075) 771-8111 (内) 3258