ロボットは学習できるか

基礎工学部　有本　卓（豊中 4505）

人間とロボットの違いの一つに、人間は学習できるが機械式ロボットは学習できない、とされていることがある。果して本当にそうなのであろうか。

ところで、現在の産業用ロボットは、機構的にも必ずしも人間の形に似せられているわけではないが、それでもあるものはだんだん人間の腕に近くなってきている（図1参照）。

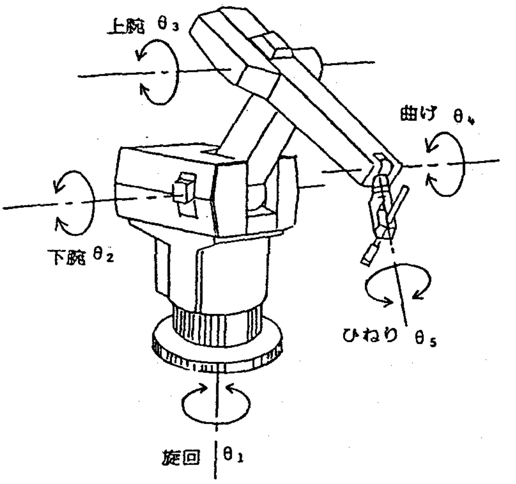

図1. 代表的な多関節型ロボット

これらのロボットは今や高度の作業をこなし、人間らしい動きをするに至っているが、現在でもそのような柔軟な動きは基本的には人間が教示することによって実現されている。すなわちオペレータがロボットの先端近くに取り付けてあるjoy−stickを握り、ロボットを動かしながら要所要所の姿勢データをコンピュータに記憶させる。これを直接教示方式という。そうでない場合も、ティーチングボックスと呼ばれる操作盤を作業員が操作し、人間の熟達した操作データをコンピュータに記憶させる間接教示方式を用いる。いずれにしても、現実の作業時には、こうして取り込んでいたデータをコンピュータにセットし、スタートさせることにより、その教示された運動を再生させる。これをティーチング・プレイバック方式、教示／再生方式という。その他、マスタースレーブ方式のマニピュレータロボットでは、人間がマスターロボットを操作することによって、人間の熟練動作をスレーブ側に伝え、精妙な作業を実現する。このように従来の機械は、たとえ現在の高級ロボットといえども、人間の教示ないし運転に依らないで自律的に動き、かつ、作業することは出来なかったのである。

人間が手や足をスムースに動かすことができるのは、生れて以来何年にもわたって自然に訓練し、運動能力を学習によって獲得してきたからである。王選手の美しい一本足打法、青木のスーパーショットも、関与の筋、腱、骨格系を操作する一連の命令プログラムが長い訓練を通じて脳のどこかに作られ、完成された上で始めて安定に実現されたのではないかと思える。実際、何らかの原因で脳に蓄積された運動のプログラムを失うと、大人といえども、長いリハビリテーションの中でプログラムを再び作り上げることによって、新たに運動機能を獲得しなければならない。このように人間は訓練と学習を通じて、

柔軟かつ精妙な運動能力を獲得したものと思われる。

さて本題に帰って、機械式ロボットは自ら学習することができるか、という問題を考えてみたい。図1のようなマニピュレータの運動方程式は、一般に、

$$R(\theta)\ddot{\theta}+f(\theta,\dot{\theta})+g(\theta)+r(\theta,\dot{\theta})=T \tag{1}$$

と表される。ここに $\theta=(\theta_1,\ \cdots,\ \theta_n)$ で、θ_i は図1に示すような関節角を表す。$R(\theta)$ は慣性行列と呼ばれ、マニピュレータの姿勢と共に変わるが、常に正定値性は保たれる。$f(\theta,\dot{\theta})$ はコリオリ力と遠心力の項であり、$g(\theta)$ はポテンシャル項に由来し、重力項と呼ばれる。これらは、たとえ複雑であっても、ロボットの各リンク要素を剛体と仮定して理論的に求めることはできる。しかし、ロボットには機械特有の摩擦、ガタ、すべり、たわみなどが起こるので、これらの測定不可能で算定できない非線形項を $r(\theta,\dot{\theta})$ と書いて外乱と称することにする。Tは一般化力であるが、図1の場合、各関節を動かすモータ(これらをactuatorと呼ぶ)の発生トルクのベクトルと考えてよい。

いま、ロボットの理想的な運動が θ の時間経過 $\theta_d(t),\ t\in[0,\ t_2]$、で与えられたとしよう。そこで各モータのサーボを操作して $T_0(t)$ というトルクを発生させ、(1)式の強制入力を $T=T_0(t)$ とすれば、ロボットの運動は(1)式の解軌道 $\theta_0(t)$ によって表されよう。ここでは前提として、(1)式の関数 R, f, g, r は判っていなくてもよいとし、ただ $\theta_0(t)$ が出力として測定できると仮定しよう。実際、各関節にポテンショメータあるいはエンコーダを取り付ければ、トルク $T_0(t)$ でロボットを起動したときの各関節の応答 $\theta_0(t)$ が得られるはずである。これは勿論、理想の運動 $\theta_d(t)$ と一致するはずはないであろう。そこで誤差ベクトル

$$e_0(t)=\theta_d(t)-\theta_0(t) \tag{2}$$

と書こう。

さて、現在ではＶＬＳＩ技術によってＩＣメモリが普及し、ロボットくらいのスピードでは、これらの関数 $\theta_d,\ \theta_0,\ e_0$ を十分高い周波数でサンプルしたデータをディジタル形式でRAMメモリに記憶させることができる。いま、これらのデータから次の入力トルク $T_1(t)$ を簡単な方式で発生させ、これでロボットを再び起動させ、応答 $\theta_1(t)$ が得られたとする。このとき問題は次の誤差 $e_1(t)=\theta_d(t)-\theta_1(t)$ が前の誤差 $e_0(t)$ よりも何らかの意味でより小さくなるような $T_1(t)$ が簡単な法則で求まるかどうかである。つまり、一般に第k回目の入力トルク $T_k(t)$ が与えられたとし、これでロボットを起動させたときの応答 $\theta_k(t)$ と誤差 $e_k(t)$ が求まったとき、これらのデータから次の入力トルク $T_{k+1}(t)$ をより良い運動が得られるように簡単な法則で決めることが出来るのだろうか。言い換えると、過去の練習データから、より理想に近いフォームを与える操作入力を自律的に発生させることが出来るかどうかが問題である。ただし、データはいつもupdateし、$\theta_d(t),\ T_k(t),\ e_k(t)$ のみから $T_{k+1}(t)$ を生成させることにしたい。また、ロボットは起動時にはいつも同じ初期条件にあり、ダイナミックスは変らないと仮定する。なお、産業用ロボットは実際には繰返しの再現精度が良くなるように設計、製作されており、この種の仮定についてはほぼ満足されている。

さて、上で述べた問題がうまく解けることが最近になって判ったのだが、これをロボットで直接見せるのは困難なので、ここでは最も簡単な例で考えてみよう。入出力関係が1次の線形微分方程式

$$\dot{y}=ay+bu \tag{3}$$

で与えられるダイナミックスについて、理想的な応答 $y_d(t)$ と第k回の入力 $u_k(t)$ と出力 $y_k(t)$ が与えられたとき、次の入力を

$$u_{k+1}(t)=u_k(t)+c\dot{e}_k(t) \tag{4}$$

と構成してみよう。これは、前回の入力に誤差の微分を適当倍したものを加えたものを次回の入力とすることを意味し、その生成はマイコンによれば容易である。このことは、まとめると、

$$\begin{aligned} \dot{y}_k(t) &= ay_k(t)+bu_k(t) \\ u_{k+1}(t) &= u_k(t)+c(\dot{y}_d(t)-\dot{y}_k(t)) \end{aligned} \tag{5}$$

が構成されることを示しており、そこで問題はこの繰返しで $e_k(t)$ が $k \longrightarrow \infty$ のとき早く0に収束するかどうか、ということになる。しかも、$e_{k+1}(t)$ は $e_k(t)$ よりも何らかの意味で（つまり、何らかの関数ノルムで）小さくなっている必要がある。ともかく、次の数値例に基づく数値計算の結果を図2に示しておく。

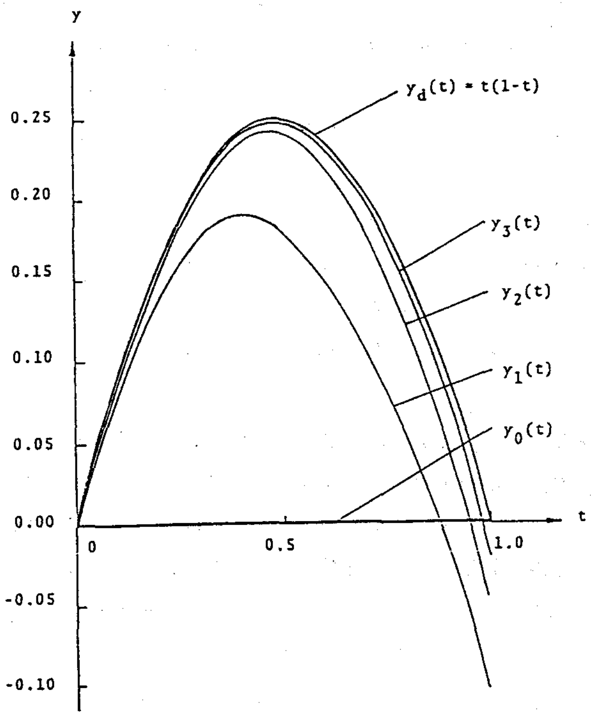

図2. " Betterment Process " の収束の様子
一次元システムの例

$a=1,\ b=1,\ c=1,$
$y_k(0)=0$
$u_0(t)=0$ for $t\in[0,1]$
$y_d(t)=t(1-t)$,

図2によると、$y_k(t)$ は急速に $y_d(t)$ に近づき、誤差 $e_k(t)$ は確実に小さくなる。

一般に、対象とするダイナミックスが

$$\begin{aligned} \dot{x} &= Ax+Bu \\ y &= Cx \end{aligned} \tag{6}$$

と表されたとき、入力を

$$u_{k+1}(t)=u_k(t)+\Gamma\frac{d}{dt}(y_d(t)-y_k(t)) \tag{7}$$

とする繰り返し過程を " Betterment Process " と名づける。理論的には、行例積CBΓがある条件を満足すれば、この " Betterment Process " は収束することが保証されるのである。この議論は更にロボットのような非線形のダイナミックスまで拡張させることができる。ロボットの場合、実際にどのように " Betterment Process " を構成するかについては触れないが、結論はこうである。「理想の運動フォームがRAMメモリに与えられたならば、機械式ロボットといえども、再現精度さえ良ければ、自律的に練習を繰り返し学習していくうちにその理想の運動フォームを実現させる能力を獲得できる。」

人間や動物の運動能力の獲得がこのような学習方式に従っているかどうかについては、専門も違うので言及するのははばかれるが、敢えて大胆に言えば、この種の案外と単純なプロセスが組まれているのではないかという想像もしている。少くとも、メカトロニクス系の制御にはこの方式が威力を発揮するのは間違いないであろう。

最後に、理想フォームの与え方の問題が残るが、これについては今のところシステマティクな方法論があるかどうか詳らかでない。