

## 制御系における状況認識と行動判断に関する研究

Study on state recognition under condition and behavior  
judgement for control systems

主任研究員：竹口知男

分担研究員：岩佐哲夫 宮井義裕 能勢和夫

本長期的共同研究は、人とインターアクションを有する移動体（自動車、自律移動の福祉ロボット等を含む）の制御系に対して、人間特性を考慮した知的処理を図るためのヒューマンインターフェイスについての研究を行うことを研究目的としている。その研究目的のために、(1)「制御系の学習機能」、(2)「制御系の状況認識と定量的表現」、(3)「自律エージェントの知識獲得」の3つの分担研究を行っている。平成11年度は「自律移動体（ロボット）の移動経路探索」のテーマの下で、分担研究(1)、(3)では「移動ロボットの移動経路に対する自律探索ならびに経路作成」について、分担研究(2)では「移動ビークルの経路誘導と障害物認識」についての研究を行い、それぞれつぎのような研究成果が得られた。

### (1) 制御系の学習機能に関する研究

行動領域内の出発点、到着点ならびに障害物の位置など行動環境についての情報が不明確な状況下で、障害物を回避移動しながらロボット自らが出発点から到達点への移動経路を探索し、学習試行を繰り返すことによって、準最適経路を自律的に探索する移動経路探索問題に対する自律移動ロボットの学習システムとして、強化学習法適用の有効性についてのシミュレーション研究を行った。シミュレーション結果として、強化学習法は自律移動ロボットの学習システムとして有効であると共に、障害物移動等による環境変化に対しても、ある環境変化の範囲内であるならば、学習効果を発揮するという結果が得られた。

### (2) 制御系の状況認識と定量的表現

移動ビークルに対して、①目的とする移動経路への誘導法と②移動経路上に不規則的に生じる障害物からの回避法が求められるため、①については誘導ラインと光センサを用いたライントレース走行を行わせる模型ビークルを作成した。また、②についてはクライアントとしてのビークルからセンシングした情報を昨年度構築した無線LANシステムによりサーバに送ると共に、サーバより他ビークルなどについてのロケーション情報や周辺環境情報を受取ることによって衝突を回避するシステムを構築中である。不規則に生じる障害物が人間である場合と他ビークル等の移動物体の場合とでは対処法が異なるため、障害物が人間か他の物体かを識別するセンサとして、通常の赤外線センサと温度差感知型赤外線センサの併用による識別システムについて検討を行った。しかし、温度差感知型赤外線センサは、センサが静止状態で設置されている場合には温度差により正常に作動するが、セ

ンサを移動体に設置すると赤外線波長が変動することから通常の赤外線センサと同じ作動状態になり、人間と他の物体とを識別する機能を果たさなかった。そのため、人間と他の物体の識別方法について再検討を行っている。

### (3) 自律エージェントの知識獲得に関する研究

移動ロボットの行動領域内の出発点、到着点ならびに障害物の位置情報が予め与えられている状況下で、最適移動経路を探索するオフライン経路作成問題に対する探索手法として、遺伝的アルゴリズム (GA) を、移動ロボットの辿れる一連のノード座標列を染色体として適用した手法を提案し、その提案手法に対する有効性をシミュレーションで検証した。シミュレーションの結果として、提案手法による解の妥当性が検証され、提案手法の有効性が確認された。

## 制御系の学習機能に関する研究

竹口知男 (工学部)

強化学習法は、設計者がロボットの動作・行動結果に対して報酬や罰 (強化信号) を設定するのみで、ロボットが自律的に行動学習を行い、不確実性環境下での行動学習を取り扱うことが可能であるといわれている。そこで平成11年度における本分担研究では、自律移動ロボットに対する学習システムとしての強化学習法の有効性について、シミュレーション研究を行うこととした。

強化学習法の有効性を検証するために、図1に示すように14×14マス目の二次元平面空間を移動ロボットの行動領域とし、任意に設定したスタート地点 (S) から出発して、途中ロボットが通過できない壁面や障害物 (O) を避け、最短ステップ数でゴール地点 (G) に向かわせる最短経路探索問題を設定した。ただし、移動ロボットには事前に障害物などの環境に関する情報は一切与えず、強化信号としては、ゴール地点到着時にそれまでの試行よりもステップ数が1ステップ減るごとに+1の報酬 (Reward) を与え、壁面や障害物に衝突するたびに-3の罰 (Penalty) を与えるようにした。また、移動ロボットは、隣接する8方向の内の一方向へある確率で1ステップずつ移動し、壁面や障害物と衝突した場合には元の座標に戻る。8方向それぞれの移動確率は、それまでの学習およびその学習試行で与えられた強化信号によってリアルタイムに変化させるようにした。

学習試行20回までのゴール地点到着に要した所要ステップ数および壁面または障害物への衝突回数の推移をそれぞれ図2および図3に示す。また、学習試行1回目、5回目、7回目および20回目における移動ロボットの移動経路を図4に示す。図2および図3より、学習が進むにつれて所要ステップ数ならびに衝突回数に改善が見られる。また、図3において、学習試行5回目から7回目において衝突回数が急増しているが、これは、図4(b)および(c)からわかるのように、壁面に近い大回りの経路からスタート地点とゴール地点との対角線方向に近い移動経路へと移動ロボット自らが変更を試みているからである。このことから、強化学習法による行動生成は、衝突回数の増加というリスクを伴いながらも

新たな行動パターンを試み、より最適な行動を獲得することも可能であることがわかる。

つぎに、行動領域上の障害物を学習試行ごとに移動させ、強化学習法が環境変化にどの程度対応できるかを検証した。環境変化の指標としては、学習試行ごとに障害物位置をランダムに移動させ、その移動個数比を環境遷移率とした。環境遷移率を1%から9%まで変化させてシミュレーション実験を行った結果、環境遷移率が3%までなら学習効果が表れたが、それ以上の環境遷移率になると学習効果を得ることができなかった。

以上のシミュレーション結果より、強化学習法は、自律移動ロボットの行動生成に有効な学習システムであり、環境遷移率が4%以下の環境変化に対しても学習効果があることがわかった。

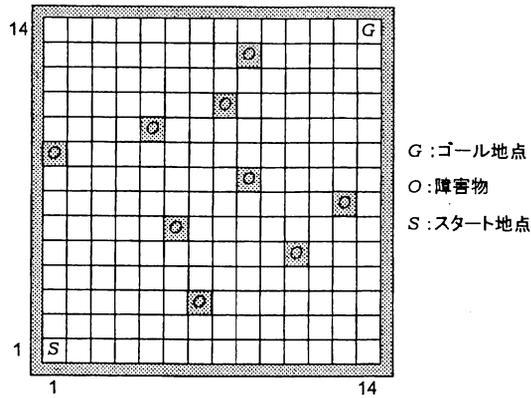


図1 ロボット行動領域

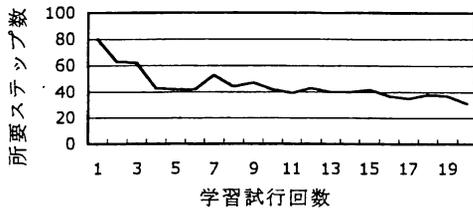


図2 所要ステップ数の推移

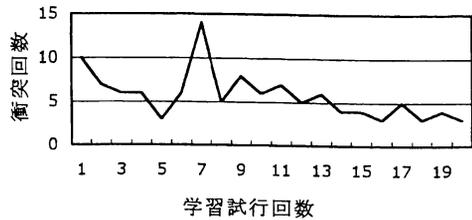


図3 衝突回数の推移

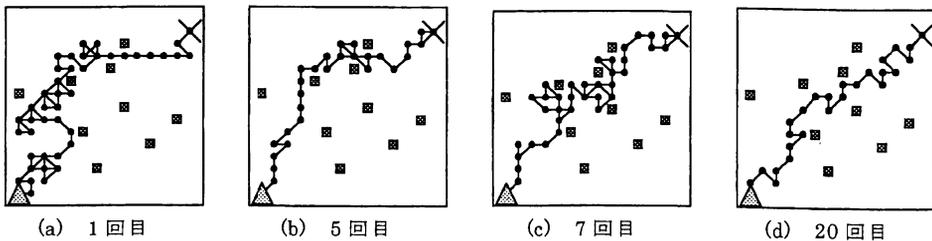


図4 学習による移動経路の推移

## 制御系の状況認識と定量的表現

岩佐哲男・宮井義裕（工学部）

移動するビークルが一定の目的を果たすためには二つの側面が要求される。その一つは目的地点に安全に到達するための手段である。もう一つはその経路において何を行うかである。前者の主題は、移動体をどのような方式で誘導するかということであつた。本テーマの延長線上に個々の移動体がネットワークのクライアントとして運行管理をするシステムの構想を持っていたために、ビークルの誘導制御は出来る限りシンプルにする事を考えていた。種々検討を行った結果、路面に一本の線（あるいはテープ）をガイドとしてビークルに取り付けた二つの光学センサによって経路から外れないように制御を行った。したがって、ビークルの進行のための制御系の状況認識の情報として路面からの明暗（0、1）を利用した。この方式による制御はセンサの取り付け幅を調節することによってビークルのふらつきを調整できる。定量的な表現とは言えないが、最もシンプルな方法であると同時に数個のセンサを追加する事によって、制御系に必要なロケーション情報の定量的な獲得が出来るメリットがあつた。その発想の起点は誘導を単純化する副産物で、従来、すべての制御を個々のビークルで行っていたのに対し、センシングした情報をサーバに送ることによりビークルの位置と次の行動をクライアントとしてのビークルが認識できることとなった。そのため、平成10年度で行った画像処理によるロケーションの把握という課題は、実用面から言えば解消できることになった。しかし、障害物の認識という課題は残る結果になった。

画像処理による周辺状況の解析は、かなり興味もあり、実現できるハード的な環境が整うまで時機を待ちたいと考えている。この障害物について、経路内に走行を妨害する障害物が存在する、あるいは、障害物が経路上にさしかかることを予測する事について検討を行った。限られた環境の中で走行するビークルの経路前方に現れる障害物は、通常、前方を走行、あるいは横断するビークルか、外乱として経路に入る人または物である。ビークルの場合は、前述のロケーションの獲得情報によって衝突を回避できる。障害物が人であると判断できれば、停止・警報・確認・発車の単純な操作のみで回避できる。物であれば、停止・サーバへの連絡・指示待ち、あるいは、安全確認・発車で回避できる。これを実現するために、温度差感知型の赤外線センサとすべての物体に反応する赤外線センサをビークルに搭載する事を検討した。しかし、人に反応するセンサは静止した状態で設置すれば、正常に作動するが、移動体に設置すると、物に反応するセンサと同じような動作状態になり、機能を果たさないことに気づいた。

原因は、赤面の思いであるが、対比する対象物が相対的に動いているため、温度差を論理的に判断できない、すなわち初歩的なミスであつた。根本的な検討とは別に、停止・警報および一定時間の待ち…を行うことによって人と物の区別を考えている。

## 自律エージェントの知識獲得に関する研究 能勢和夫（工学部）

2年目は、遺伝的アルゴリズムによるロボット移動経路の作成について研究した。すなわち、2次元平面内の与えられた領域において、出発点、到着点、障害物の位置情報が予め与えられたとき、ロボットの最適移動ルートを移動開始前に作成するオフライン経路作成問題について、遺伝的アルゴリズム適用の観点から検討した。

移動経路の作成とは、出発点、到着点、障害物の位置情報に基づいて折れ線の一連のノード座標を求めることである。遺伝的アルゴリズムでは、移動経路の候補を個体の染色体として表現する。ここでは、染色体はロボットが辿る折れ線の一連のノード座標列とした。特に、染色体の長さは個体ごとに可変とし、その上限は障害物の数から設定した。障害物は壁形状とした。また、経路が実行不可能であるとは、その経路が障害物と交わる場合をいう。

考案した遺伝的アルゴリズムの手順を図1に示す。図中の各項目を①～⑨に説明する。

- ①初期個体群の生成：初期個体における各ノード位置は、与えられた2次元平面の矩形領域内において確率的に定める
- ②適合度：経路評価は、経路を構成する各線分が障害物と交差しているか否かのチェック、経路の評価項目算出などである。各染色体に対して適合度を評価値の逆数で設定する。
- ③停止条件：最大世代数を定め、遺伝的操作の繰り返しを停止する。
- ④操作選択：遺伝的アルゴリズム適用のポイントは、操作選択の導入である。これは、交叉、突然変異、挿入あるいは削除の操作を確率的に選択するものである。その際、適合度向上に効果的な操作を高い確率で適応的に選択しようとするのが操作選択である。
- ⑤交叉：個体群の中からルーレット選択により選出した2個の親個体の染色体に対して、一点交叉を適用し、子個体を2個作り出す。その際、2個の親個体の染色体長さが異なること、染色体長さに上限があることを考慮して、各親個体の交叉位置を確率的に決める。
- ⑥突然変異：個体群の中からルーレット選択により選出した1個の親個体の染色体に対して、経路が実行可能経路である場合には染色体の各遺伝子に対して突然変異を行なうか否かを確率的に定め、突然変異を施す場合には小変位突然変異操作を行なう。一方、実行不可能経路に対しては、障害物を迂回するために障害物と交差する線分のノード座標を大きく変化させる大変位突然変異を行なう。
- ⑦挿入：個体群の中からルーレット選択により選出した1個の親個体の染色体に対して、経路が実行可能である場合には、すべてのノード間を対象に挿入するノード位置を確率的に定める。実行不可能経路に対しては、障害物と交差する線分のノード間に新たなノードを追加する。なお、交差する線分が複数の場合は確率的に一つを選ぶ。いずれの場合でも、挿入ノードの座標は確率的に定める。
- ⑧削除：個体群の中からルーレット選択により選出した1個の親個体の染色体に対して、経路が実行可能である場合には、平坦な経路はできるだけ直線となるように、経路平滑

度に応じて削除するノードを確率的に定める。実行不可能経路に対しては、障害物と交する線分のノードから確率的に一つを選び、削除する。

⑨世代交代：世代交代は2通りの方法を準備した。すなわち、親個体と子個体を併せた集団の中で適合度の良い個体を次世代の親として残す方法と、生成した子個体全部と親個体群の中で適合度の良いものを次世代の親として残す方法とである。

提案手法の有効性を評価するために数値実験を行った。一例として図2に示す配置に関して最適経路を探索した結果を図3に、各遺伝的操作の選択確率の推移を図4に示す。図3から解の妥当性が、図4から世代の経過に従って選択確率が変化していく様子が分かる。このような数値実験を通して提案手法の妥当性が確認できた。

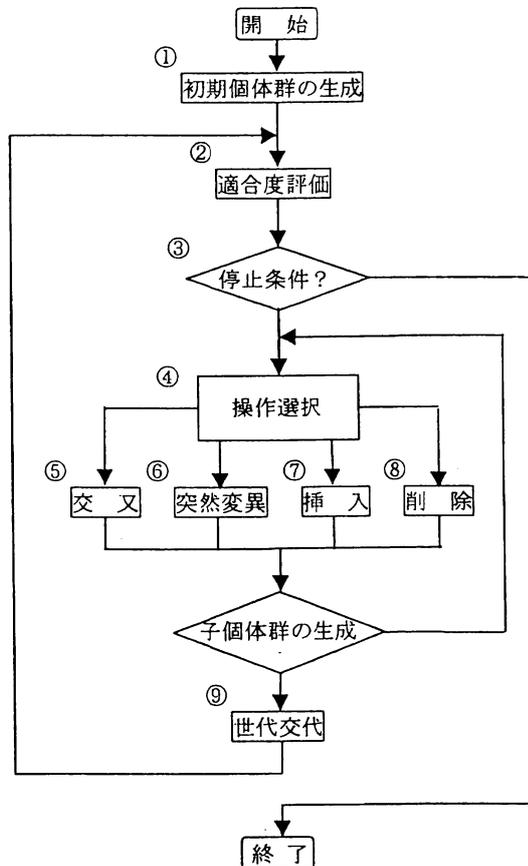


図1 適応的遺伝的アルゴリズム

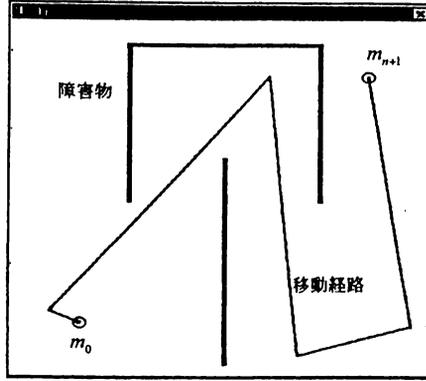


図2 探索平面の配置

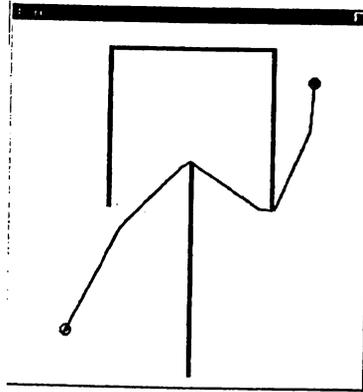


図3 探索結果

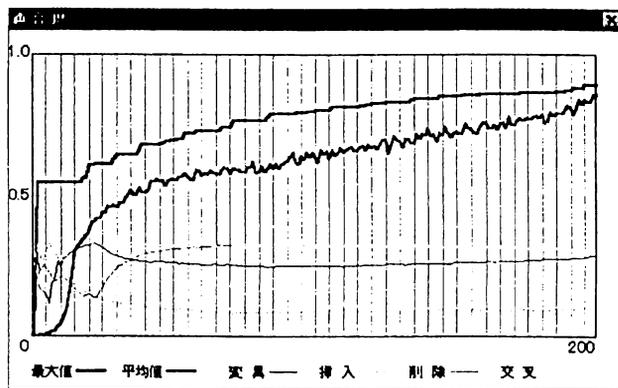


図4 選択確率の推移