遺伝的アルゴリズムによる仮想的な移動ロボットの 環境変化に対応する一般化知識の獲得手法

非会員 大 野 宏 司 (豊田中央研)
正 員 古 橋 武 (名古屋大)
正 員 内 川 嘉 樹 (名古屋大)

A Method of the Generalized Skill Acquisition of Virtual Mobile Robot for Adapting Environmental Changes using Genetic Algorithms

Hiroshi Ohno, Non-member (Toyota Central Research & Development Labs., Inc), Takeshi Furuhashi,
Member, Yoshiki Uchikawa, Member (Nagoya University)

We propose a method for acquisition of generalized action-decision rules for mobile robot which is based on the assumption that selection of adequate focal point is important for the skill acquisition. The method makes it effective to acquire the generalized knowledge which are robust against environmental changes, especially unknown environment. The skill acquisition starts from discovery of rules to drive a robot to a goal point under a certain environment using genetic algorithm. Then, for the generalization of acquired rules, we select an adequate sensing range of the robot for detecting obstacles, which is described in the if-part of rules. Results of computer simulation show that the extracted skills work well to accelerate the learning of robot under new environment.

キーワード：スキル獲得，移動ロボット，遺伝的アルゴリズム，一般化知識

1. はじめに

ピアノ演奏や車の運転などのスキル獲得では、目の付け所（注視点）をいかに選択するかが習熟度の向上において非常に重要である。そこでは、行動の習熟度により注視点の選択に違いが見られる。例えば、三浦、高山、三木⁽¹⁾らは技能水準の高いピアノ演奏者の読譜の範囲が、技能水準の低い演奏者に比べて広いことを示している。そこで、移動ロボットやエキスパートシステムなどの知的システムにおける知識獲得の問題を取り上げ、注視点と習熟度の関係に着目した検討を行う。

従来研究において注視点と知識獲得または習熟度の関係を考察したものはいくつかみられる。榎木、片井、岩井⁽²⁾らは、オペレータの操作履歴の解析を行い、操船履歴の定性的解釈から注視点を決定している。そこでは、オペレータが注視点を目標として操船のサブプランの設定を行っていることが示されている。また、塩瀬ら⁽³⁾⁽⁴⁾は、移動ロボットを題材として、熟達と注視点の形成について検討を加えている。そこでは、進化的計算の過程で環境が提供する情報の変遷により、アフォーダンス⁽⁵⁾と注視点の変容とを関係付けて論じている。更に、文献(4)では習熟において環境を重要視する立場でスキルの転移をアフォーダンス

に関連付けて論じている。

一般に知的システムの学習は特定の環境下で行われる場合が多い。そのため、環境変化、特に未知環境に対するロバスト性が低いことが問題となる。

一方、人の熟達者（適応的な熟達者と呼ばれる）は、経験を通して知識の一般化を行い課題状況の変化に柔軟に対応することができる。そこで、本研究では、適応的な熟達者にみられる特徴を参考にして、注視点と習熟度の関係およびスキルの転移による環境変化、特に未知環境に対するロバスト性の向上について検討を行う。ここで、未知環境を取り上げたのは、環境変化において最も厳しい条件であると考えたからである。スキルの転移による環境変化に対するロバスト性については、適応的な熟達者にみられる特徴の一つであり⁽⁶⁾、知的システムの学習において重要な要件であると考えられる。そこで、仮想的な移動ロボットを題材として、遺伝的アルゴリズム（以下、GA）により行動判断ルールの学習（スキル獲得）を取り上げる。そして、ロバスト性を高めるためには、獲得されたスキルを環境に依存しない一般化した知識として利用することが必要であると考えられる。そのために、注視点に着目した一般化知識をスキルから獲得するアルゴリズムを提案する。この一般化知識を学習時とは異なる環境下における学習に適

用することで、スキルの転移による環境変化（未知環境）に対するロバスト性の向上が実現できることを示す。この結果は、また、適応的な熟達者の特徴を採り入れた知的システムの実現の可能性を示唆する。

一方、塩瀬ら⁽⁴⁾の研究ではアフォーダンスの考えを援用してスキル転移を注視点に着目した観点から検討を加えている。ただし、アフォーダンスを工学的に利用する立場からは、環境のアフォーダンスを陽に特定（記述）し、スキル転移の議論を進める必要があると考えられる。これに対して、著者らのアプローチでは注視点の選択により行動の変化をもたらすようにルールを記述し、獲得されるルールの一般化によりスキル転移を可能にする。

2. 提案手法

ここでは、移動ロボットの構成、GA による行動判断ルールの学習および提案する一般化知識獲得アルゴリズムについて述べる。

〈2・1〉 移動ロボットの構成 移動ロボットはその身体の大きさはなく、進行方向に等角度に分けられた五つの領域を対象とするセンサを備えたものである（図 1 参照）。センサの機能は障害物の有無を特定する。また、センサの各領域は距離により三つの領域（NEAR, MIDDLE,

図 1 移動ロボットのセンサ領域
Fig. 1. Sensing ranges of mobile robot.

図 2 移動ロボットの行動判断用プロダクションシステム
Fig. 2. Block diagram of production system.

FAR）に分けられている。ロボットに対する行動指令は、前進（MOVE-FORWARD）、停止（STOP）、左旋回（TURN-LEFT）、右旋回（TURN-RIGHT）および継続（CONTINUE）の五つである。前進は速度一定で進むことを行う。左右旋回は速度一定で 90 度進行方向を変えることを行う。継続は次時刻の行動指令が現在のものと同じであることを意味する。速度の大きさは単位時間あたりの単位長さとした。ロボットの行動判断はプロダクションシステムにより行われる（図 2 参照）。IF-THEN ルール（以下、ルール）の競合解消は等確率で選択することにより行う。ルールの形式は、

IF(action $\wedge$ sensor $\wedge$ t/f $\wedge$ range) THEN (action)
とした。IF 部では、現在の行動指令 (action) とセンサの番号 (sensor)、障害物の有無を示すフラグ (t/f) およびその領域 (range) の論理積の真偽により発火するかどうか決められる。THEN 部では、次時刻の行動指令が指定される。IF 部に現在の行動指令を入れた理由は、ルールの競合を減らす働きと、後述する一般化知識の獲得のためである。例えば、以下のルールの意味は、

IF (STOP $\wedge$ 3 $\wedge$ FALSE $\wedge$ NEAR) THEN
(MOVE-FORWARD)

現在の行動指令が STOP で、かつセンサ 3 の領域 NEAR に障害物がない (FALSE) ならば、MOVE-FORWARD を行動指令とすることである。

〈2・2〉 行動判断ルールの学習 移動ロボットに与えられる目的は仮想的な環境内に設定されたゴールに入ることである。そのために必要な一連のルールを GA により学習する。図 3 に示す手続きにより、移動ロボットの行動判断ルールの学習が行われる。

次に、染色体へのルールのコーディングの仕方を表 1 に示す。一つのルールは 12 ビットで表現される。一つの染色体には複数のルールがコーディングされる。また、センサについては前方 (3) と近接 (NEAR) の出現頻度が高くなるようにバイアスが掛けられている。これは、行動において一般に進行方向と自分の近くをよく見るのが自然であると考えたからである。

次に、設計者がロボットに与える習熟の評価を「なるべく早くゴールに到達する」とこととした。そして、習熟の評価を反映するように適応度関数 (fitness) の定義を以下のようにした。

$$\text{fitness} = \begin{cases} 1/\text{time} & \text{success} \\ 0 & \text{fail} \end{cases} \dots\dots\dots (1)$$

ここで、time はゴールに到達したときの走行時間である。success (成功) はゴールに入った場合（ゴールで停止する必要はない）、fail (失敗) はある一定時間動かなかった場合とゴールに到達せず障害物に衝突した場合である。また、GA のオペレータは一点交差およびエリート保存によるランキング戦略とした。

〈2・3〉 一般化知識獲得アルゴリズム 適応的な熟達者の特徴である環境変化に対する高いロバスト性を実現する

表 1 ルールのコーディング
Table 1. Rule coding method.

action	bit patrn (3 bit)
MOVE FORWARD	000, 001
TURN-LEFT	010, 011
TURN-RIGHT	100, 101
STOP	110
CONTINUE	111
sensor	bit pattern (3 bit)
1	000
2	001
3	100, 101, 110, 111
4	010
5	011
t/f	bit pattern (1 bit)
FALSE	0
TRUE	1
range	bit pattern (2 bit)
NEAR	00, 11
MIDDLE	01
FAR	10

図 3 移動ロボットの行動判断ルールの学習手続き

Fig. 3. Procedure for learning action-decision rules.

ためには、獲得した知識を何らかの仕方一般化し適用することが必要であると考えられる。すなわち、環境に依存

Step 1) Rule extraction

- 1-1) find rules which produce action changes in skillful rules
- 1-2) find rules which maximize fitness value when those rules replace the corresponding rules in initial rules

Step 2) Rule generalization

extract the range from IF-part of the rule in Step 1)

図 4 一般化知識獲得アルゴリズム

Fig. 4. Algorithm of the acquiring method for the generalized knowledge.

しない知識をもつ必要がある。そこで、注視点に着目した一般化知識獲得アルゴリズムを図 4 に示す。

アルゴリズムは染色体からデコードされたルールを対象にして、ルールの抽出と一般化の手続きからなる。ルール抽出では、始めに、習熟ルールから行動に変化をもたらすルールを抽出する。これは、習熟度の向上においては、行動に変化をもたらすルールが重要であると考えたからである（知識獲得における仮定）。そして、その中から IF 部と THEN 部の action がそれぞれ等しい初期ルール中の対応するルールと交換し適応度（習熟度）を求める。次に、抽出されたルールの中で、適応度の（初期ルールに比べ）増加が最も大きくなったルール（抽出ルールと呼ぶ）を選択する。ここで、初期ルールとは適応度（習熟度）の低いルール群（ゴールに到達することが出来る）である。これは設計者が学習時に初期値として与える。これは初心者が専門家からアドバイスを与えられることに相当する。行動の変化は、ルールの IF 部と THEN 部の action の違いにより調べる。ただし、適応度（習熟度）の向上には複数のルールが関与すると考えられる。今回は、抽出ルールは単独（適応度の向上を最も高めたルール）であるとした。

また、注視点に着目した一般化を行うために、抽出ルールから、ルールの IF 部における range の部分だけを抽出して一般化知識とする。これは、センサの番号や障害物の有無に関する知識は環境に依存する所が大きいと考えられ、また注視点に着目する点からも range の部分だけを対象とした。また、このように獲得された一般化知識は、習熟度の評価である「なるべく早くゴールに到達する」という基準のもとに適用される。

次に、一般化知識を学習時に適用するために以下のようにした。ただし、行動に変化をもたらすルールであるため、IF 部と THEN 部の action が違うルールの場合にのみ一般化知識は適用される。

- 適応度が 0 より大きくなった時点でエリート保存された染色体に一般化知識を適用し染色体（各ルール）を変更する

ここで、適用を適応度が 0 より大きくなった時点にしたことおよびエリート保存された染色体に限定したことの理由は、学習の初期における多様性を維持するためである。

3. シミュレーション実験および結果

〈3・1〉 環境と課題 移動ロボットが動き回る環境を図5に示す。図中ハッチが掛けられた部分が障害物である。ロボットに与えられる課題はスタート地点からゴールに到達することである。ここで、スタート地点は右下の通路中央であり、ゴールは左上のライン上の地点である。図中の座標値の単位は単位長さである。

〈3・2〉 行動判断ルールの学習 ここでは、習熟度により注視点がどのように変化するかを調べる。初期ルールを表2に示す。初期ルールが一つの染色体にコーディングされる。一つの染色体の長さは四つのルールからなるため48ビットである。これらのルールはロボットが自分の近接の障害物に注意しながらスタート地点からゴールに到達するためのルールである。

学習条件は、世代数100、染色体数30、ランキング戦略における染色体の選択数10、crossover rateおよびmutation rateはそれぞれ0.9、0.6とした。ここで、30の染色体にはそれぞれ初期ルールがコーディングされている。学習結果について、最良個体の適応度（習熟度）と世代数の関係を図6に示す。

図6より6世代以降、適応度の増加は見られなかった。ただし、この結果はあるランダムシードで得られたものであり、別のランダムシードでは学習は遅くなる場合があった。100世代目の最良個体の染色体をデコードして得られるルール（習熟ルール）を表3に示す。

図5 移動ロボットの環境
Fig. 5. Simulation environment.

表2 初期ルール
Table 2. Initial rules.

No.	action	sensor	t/f	range	action
1	STOP	3	FALSE	NEAR	MOVE-FORWARD
2	MOVE-FORWARD	3	FALSE	NEAR	CONTINUE
3	MOVE-FORWARD	3	TRUE	NEAR	TURN-LEFT
4	TURN-LEFT	3	FALSE	NEAR	MOVE-FORWARD

図6 学習結果—適応度と世代数の関係—
Fig. 6. Simulation result: Profile of the fitness function.

図7 移動ロボットの走行軌跡
Fig. 7. Running trajectories.

表3 習熟ルール
Table 3. Skillful rules.

No.	action	sensor	t/f	range	action
1	STOP	3	FALSE	NEAR	MOVE-FORWARD
2	MOVE-FORWARD	3	FALSE	FAR	CONTINUE
3	MOVE-FORWARD	3	TRUE	FAR	TURN-LEFT
4	TURN-LEFT	3	FALSE	NEAR	MOVE-FORWARD

また、初期ルールと習熟ルールによる移動ロボットの走行軌跡を図7に示す。

初期ルールと習熟ルールの比較から、習熟度の向上をもたらす原因はセンサの領域を NEAR から FAR にすることであることがわかる。すなわち、早くゴールに到達するという習熟の評価のもとで、注視点を遠くに置き、障害物を早くとらえることが習熟度を高くすることになる。

以上のシミュレーション実験から、適切な注視点の選択が習熟度の向上につながることを移動ロボットの行動判断ルールの獲得を題材にして示すことができた。ただし、習熟度の向上に着目しているため、今回のシミュレーション実験から得られた習熟ルール（表3参照）が最適な、すなわち最大の適応度をもつルールである必要はなかった。

〈3・3〉 一般化知識の効果 ここでは、獲得されたルールから一般化知識を獲得し、学習とは異なる条件下で行動判断ルールの獲得に適用し習熟促進効果を調べる。そして、習熟促進効果からスキル転移による環境変化（未知環境）に対するロバスト性を判断する。

図4に示した一般化知識獲得アルゴリズムにより一般化知識を獲得する。抽出ルールは、

IF (MOVE-FORWARD $\wedge$ 3 $\wedge$ TRUE $\wedge$ FAR)
THEN (TURN-LEFT)

であった。次に、これを注視点の観点から一般化すると、一般化知識は、

IF ($x_1 \wedge x_2 \wedge x_3 \wedge$ FAR) THEN (x_4)

となった。ここで、 $x_1 \sim x_4$ は不定であり何にでも整合することを示す。ただし、IF 部と THEN 部の x_1, x_4 は両

者の action が異なるものであれば何でもよいことを示す。

次に、学習時の環境が変化することによる最も厳しい条件として、ロボットが環境に関する知識を何ももたない状態を想定するために（これは、ロボットが未知環境に置かれたことになる）、初期ルールがすべてランダムに、すなわち各染色体がランダムに設定される条件で、習熟度の促進効果が見られるかを調べる。

学習条件は、世代数 100、染色体数 30、ランキング戦略における染色体の選択数 10、crossover rate および mutation rate はそれぞれ 0.9、0.6 とした。また、一つの染色体にコーディングされるルール数を六つとした。

16種のランダムシードでシミュレーションを行い、13種の場合（8割強）で習熟促進効果が見られた。世代数に関する最良個体の適応度の変化の一例を図8に示す。図には一般化知識を適用した場合（実線）としない場合（破線）が示されている。適応度は一般化知識を適用することで改善されている。次に、促進効果が見られなかった場合を図9に示す（最良個体の場合）。この場合では、ランダムに設定された初期ルールの組合せが悪く一般化知識をいかなることができない。これはプロダクションシステムのルールの競合解消を等確率で行っていることに起因していると考えられる。また、促進効果が見られないほかの例の原因としては、ランダムシードの設定によっては一般化知識を適用することにより染色体の多様性が低下してしまうのではないかと考えられる。

以上のシミュレーション実験から、一般化知識の適用により学習時とは異なる環境（未知環境）下でも習熟度の促

図8 一般化知識の効果検証（効果ありの場合）
Fig.8. Simulation results for testing the validity of the generalized knowledge (effective case).

図9 一般化知識の効果検証（効果なしの場合）
Fig.9. Simulation results for testing the validity of the generalized knowledge (ineffective case).

進効果が見られた。これは、適応的熟達者が示す特徴の一つを仮想的な移動ロボットの行動判断ルールの学習を題材にして示したことであり、知的システムにおいて注視点に着目した一般化知識獲得アルゴリズムの重要性を示唆するのではないかと考えられる。

4. ま と め

本研究では、習熟度と注視点の関係に着目し、仮想的な移動ロボットの行動判断ルールの獲得の実験を通して、適切な注視点の選択が習熟度の向上に重要であることを示した。また、適応的な熟達者にみられる特徴の一つである環境変化、特に未知環境に対するロバスト性の向上に関して、提案した一般化知識獲得アルゴリズムにより可能であることを示唆する結果を得た。すなわち、スキルの転移をルールの視点からとらえた注視点に着目する一般化知識の有効性を示した。

本手法の課題としては、一般化知識が単独のルールで表現されるため、複雑な判断を伴うような場面には適用できないことや、一般化知識の効率的な獲得の仕方があげられる。

更に今後の研究としては、行動と環境の相互作用の過程で見いだされた（行動を決定した）情報の一般化の可能性やそれと習熟との関係について調べるのが重要であると思われる。

(平成9年2月17日受付, 同9年8月26日再受付)

文 献

- (1) 三浦利章・高山典子・三木睦子:「ピアノ視奏時の情報獲得・処理: 曲の難易度と技能水準を中心として」, 音楽研究, 6, 9~40(昭63)
- (2) 榎本哲夫・片井 修・岩井壮介:「オペレータの操作履歴からの制御則の自動チューニングと知的ファジィ制御」, 計測自動制御学会論文集, 26-8, 854~861(平2)
- (3) 塩瀬隆之・榎本哲夫・山下剛史・片井 修:「進化型計算によるアフォーダンス概念の獲得と非言語コミュニケーションに関する

考察」, 計測自動制御学会第23回知能システムシンポジウム講演論文集, 53~58(平8)

- (4) 塩瀬隆之・榎本哲夫・山下剛史・片井 修:「自律移動体の行動からの創発的注視点形成」, 第12回ファジィシステムシンポジウム講演論文集, 899~902(平8)
- (5) 佐々木正人:アフォーダンスー新しい認知の理論, (平6)岩波書店
- (6) 大浦容子(波多野(編)):熟達化, 認知心理学5学習と発達, (平8)東京大学出版会

大 野 宏 司 (非会員) 1984年岐阜大学工学部卒業。86年同大大学院工学研究科修士課程修了。88年より(株)豊田中央研究所入所。現在、知的システムの学習およびドライバのモデル化の研究に従事。

古 橋 武 (正員) 名古屋大学大学院工学研究科電子情報工学専攻。1985年3月同大学院博士課程修了。90年12月より同大工学部助教授、現在に至る。工学博士。94年日本ファジィ学会論文賞受賞。主として、ソフトコンピューティングに関する研究に従事。

内 川 嘉 樹 (正員) 名古屋大学大学院工学研究科電子情報工学専攻。1969年同大学院博士課程満了。85年同大工学部教授、現在に至る。工学博士。ニューラルネットワーク、ファジィ理論、遺伝的アルゴリズム、自律分散システム、バーチャルリアリティー、電子計数法などを利用した、システム技術の基礎および産業への応用に関する研究に従事。