

既存知識の統合による複合的知識獲得モデル

Combined Knowledge Acquisition Model by Integration of Existing Knowledge

矢部 達也[†]
Tatsuya Yabe服部 元信[‡]
Motonobu Hattori

1. はじめに

人間が新たなスキルを習得するとき、まずはそれに関する規則や定石などを教わり、それらに忠実に従った、いわば教科書どおりの行動していると考えられる。これはそのときに必要な規則を適用しながら行う意識的な行動の過程といえる。しかし、ある程度同じ行動をしていると、次第に規則や個々の行動などを気にせずとも無意識に同様の行動が取れるようになってくる。例えば、サッカーのドリブルを考えてみると、初心者にはドリブルをするとき、ボールを蹴っては向かって行くという2つの行動をただ繰り返す。だが、そのように走っては蹴ってを意識的に繰り返しているうちに、いつの間にかそのようなことを意識しなくてもドリブルができるようになっていくだろう。

無意識的な行動の過程は、環境からの情報が与えられるとすぐさま行動が得られる過程である。これはある一つのニューラルネットワークに入力が与えられると、すぐさま出力が得られる過程として考えられる。本稿ではこの過程を知識と呼ぶこととする。さらに意識的な行動の過程とは、既存の知識を if-then 的な規則により用いながら行動する過程であると考えられる。そうすると意識的な過程からの無意識的な行動の形成とは、if-then 文的な規則による行動の繰り返しのなかで並列的に、あるニューラルネットワークがその行動を手本として学習されているのではないかと考えられる [1]。

本研究では、規則的な行動の繰り返しにより、既存知識を統合した複合的知識を獲得するモデルを提案する。これにより、同様の知識を既存の知識を用いず 0 から学習させた場合と比べ、学習時間が短縮され効率的に知識を得ることができると考えられる。そこで提案モデルの有効性を検証するために、仮想ロボットによるシミュレーション実験を行った。

2. 既存知識の統合による複合的知識獲得モデル

前述のように、知識を学習済みのニューラルネットワークであると考え、意識的な行動の過程とは、複数の既存知識とその if-then 的な規則を、新たに一つの知識として統合することと考えることができる。そこで、本研究ではそのしくみを模したものととして図1のモデルを提案する。

点線枠内の処理では、まず環境の状態が入り、既存知識 1~N のニューラルネットワークのどれを用いる

かが、ある if-then 的な規則に従って選択される。そして、選択されたニューラルネットワークに環境の状態が入力され、その出力により行動が決定される。さらにそれと並行して、新たな知識 X のニューラルネットワークにも同様の環境の状態が入力され、その教師としていま選択された既存知識のニューラルネットワークによる出力が与えられ、誤差逆伝搬 (Back Propagation:BP) 法により学習を行う。こうして何度も同様の行動をこなすことで、知識 X のニューラルネットワークは個々の既存知識による行動と選択規則について反復的に学習することとなり、最終的に、目的行動を取れるように学習された一つのニューラルネットワークが構築される。このようにして得た新たな知識は、複数の既存知識を取り込んだ複合的知識であるといえる。

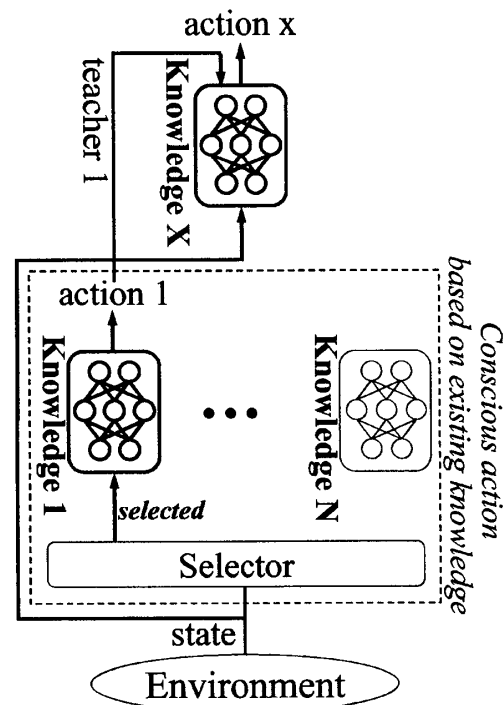


図1: 既存知識の統合による複合的知識獲得モデル

3. 0からの知識の獲得

複合的知識を獲得するにあたっては、いくつかの基礎的な知識を 0 から獲得する必要があるが、その獲得には Direct-Vision-Based 強化学習 [2] を用いる。また、比較対象となる複合的な知識の獲得にもこの方法を用いる。

[†]山梨大学大学院医学工学総合教育部
[‡]山梨大学大学院医学工学総合研究部

強化学習は、報酬と罰という少ない情報から試行錯誤的に行動を獲得していくという、自律性・柔軟性の高さから注目を集めている。また、ニューラルネットワークにおいては、その有用な学習法としてBP法があるが、これには教師信号が必要であり、人間が訓練データを用意してやる必要がある。場合にもよるが十分な量・質の訓練データの確保には困難が伴う場合が多い。そして、その二つを融合した手法である Direct-Vision-Based 強化学習は、センサからモータまでのシステム全体の制御をニューラルネットワークにより構成し、強化学習に基づいて教師信号を生成し学習を行うことで、自律的・柔軟的な学習を実現している。

4 計算機シミュレーション

ロボットシミュレータ Webots を用いて、移動ロボットに提案モデルを実装し、計算機シミュレーションを行った。

まず、視界に目標物があればそれへと向かって進む『目標物到達』、目標物が視界になければ視界に納めようと探索する『目標物探索』という2つの基礎知識として既に獲得しているものとする。その上で、「目標物が視界にあれば『目標物到達』、なければ『目標物探索』を選択」という規則により使い分けることで、目標物を視界に捉えるよう探索し、さらにそこへ向かっていくという行動をする複合的知識『目標物探索到達』を獲得するタスクについて実験を行う。そのイメージを図2に示す。

また、比較実験である同様の知識の0からの獲得は、まっさらの状態からの学習である。そのため最初から遠くの目標物に辿り着くのは極めて困難であり、学習に必要以上の時間がかかってしまう。そこで、最初は目標物をロボットの近い位置に配置し、学習の進行に従って少しずつ離していくという方法を使った。そのため、比較する対象である提案モデルの場合もそれに合わせて同様の方法を採用した。しかし、提案モデルは完成した個々の知識を用いるので、そのような方法を使わずに目標物の配置の自由度を最大にした状態から始めても学習が可能である。

4.1 既存知識の統合による複合的知識の獲得

基礎知識である『目標物到達』と『目標物探索』は、事前に Direct-Vision-Based 強化学習により獲得しておく。その2つの知識と、前述の規則を用い、提案モデルにより2つの知識の統合した。

目標物に到達するか、見失うか、規定時間に達するかを1試行とし、新知識のネットワークの学習度を調べるため、100試行毎に新知識のネットワークのテストを500回行った。そして、その成功率が95%を越えた場合には、目標物の配置範囲を少し拡大し、学習の段階をステップアップさせた。また、目標物の配置範囲が予め定めた最大値を越えたときを学習の完了とした。比較実験である0からの知識獲得においても同様の方

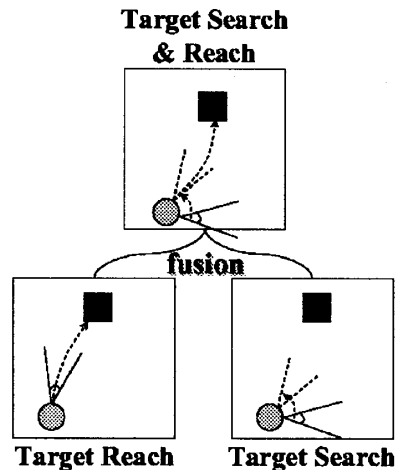


図2: 目標物探索到達タスク

法を用いた。2つの実験の結果を表1に示す。提案モデルを用いて既存知識を統合した場合の方が、Direct-Vision-Based 強化学習により0から学習した場合よりも約半分程度の学習時間で知識を獲得できていることが分かる。

なお、既存知識の統合において、始めから配置範囲を最大にして学習させた場合、学習時間は10回のシミュレーションにおいて平均650試行で済んだ。さらにこのとき、学習時の規則的な行動とは異なる、『目標物到達』と『目標物探索』を規則による切替の境界付近で融合したような、継目のないシームレスな行動を獲得するケースが多く見られた。

表1: 複合的知識『目標物探索到達』の獲得結果

	既存知識の統合	0からの学習
学習時間	2520 試行	4450 試行

5. まとめ

規則的な行動の繰り返しにより、既存知識を統合した複合的知識を獲得するモデルを提案した。その結果、提案モデルにより知識を統合することで、0から知識を得るよりも効率的に獲得できることが示された。また、提案モデルを用いることによって、意識的な行動から、規則と既存知識が融合された無意識的ともとれるような自然な行動を獲得することができた。

参考文献

- [1] 信原幸弘. “考える脳・考えない脳一心と知識の哲学”, 講談社, 2000.
- [2] 柴田克成, 岡部洋一, 伊藤宏司. “ニューラルネットワークを用いた Direct-Vision-Based 強化学習 – センサからモータまで”, 計測自動制御学会論文集, Vol.37, No.2, pp.168-177, 2001.