

インタラクションを通じた学習

麻生 英樹 (h.asoh@aist.go.jp)

産業技術総合研究所

情報処理研究部門 メディアインタラクショングループ

機械学習の分野では、これまで主として、事例からの学習についての研究が活発に行われてきた。そこでは、多くの場合、あらかじめ大規模なデータが用意され、学習システムはそれを用いてオフラインで学習する。しかしながら、知能ロボットに代表されるような知的なシステムが日常的な生活の場へと進出するに伴い、よりインタラクティブな、使用される現場での学習能力がシステムに求められるようになってきている。さらに、学習対象となる知識が複雑な構造をもつものになっていることもまた、こうしたインタラクションを通じた学習の必要性を高めていると言える。本稿では、このような、インタラクション指向の学習、インタラクションを通じた学習に関する研究について、これまでの事例を概観しつつ、今後の可能性と課題について考察する。

An Essay on Learning through Interaction

Hideki Asoh (h.asoh@aist.go.jp)

Media Interaction Group

Information Technology Research Institute

National Institute of Advanced Industrial Science and Technology (AIST)

In the field of traditional machine learning, many researches focused on learning from examples. In most of cases the learning systems learn from prepared large amount of data in off-line mode. As various intelligent systems, such as intelligent robots, are becoming to enter into our everyday life recently, however, more interactive, on-the-fly learning capability is required for the systems. In addition, the fact that the knowledge which the systems should acquire is becoming complex one with combinatorial structures, also makes the needs for interactive learning scheme increase. In this report, we will overview the various approaches to learning through interaction, or interaction-oriented learning, and investigate the possibility of the research direction and future research issues.

1. はじめに

機械学習[1]の分野では、これまで主として、事例からの学習についての研究が活発に行われてきた。そこでは、多くの場合、あらかじめ大規模なデータが用意され、学習システムはそれを用いてオフラインで学習する。このような学習アルゴリズムは、パターン認識の分野で大

きな成功を収めている。機械学習研究の現在の課題は、大きく分けて、より大規模なデータへの対処と、より複雑でダイナミックな対象への対処の二つである。前者は、いわゆるデータマイニングと呼ばれる研究分野に象徴される。一方、知的なシステムが、工場などのあらかじめ限定され、環境が固定された世界から、実世界、

日常的な生活の場へと進出するに伴い、より複雑でダイナミックな学習の必要性が高まっている。そこにおいては、インタラクティブな、使用される現場での学習能力がシステムに求められるようになってきている。学習対象となる知識が複雑な構造をもつものになっていることもまた、こうしたインタラクションを通じた学習の必要性を高めている。このようなインタラクションを通じた学習は、主に知能ロボットの分野を中心に発展してきたが、それ以外にも、いろいろな分野で比較的独立に行われてきている。本稿は、こうしたインタラクションを通じた学習の研究について、これまでの試みを鳥瞰的な視点から検討することを目的とする。そのために、まず、機械学習や認知科学の研究の文脈でこれまでに提出されてきたインタラクティブ指向な学習スキームの中から、基本的なもの、あるいは興味深いものとして、能動学習、強化学習、模倣学習、MAT 学習、進化的アプローチ、状況論アプローチ、を取り上げ、概観する。次に、著者らによる対話にもとづく学習の事例について少し詳しく検討する。さらに、それらの学習におけるインタラクションのあり方を整理する。最後に、今後の可能性や課題について考察する。

2. インタラクション指向の学習研究

2.1. 能動学習

通常の事例からの学習では、学習用データを人手で収集するが、システム自身が能動的に質問を作成し、それに対する回答を得ることで学習用データを集めるのが能動学習である[2]。これによって、学習結果に重要な影響を与えるクリティカルなデータを選びながら収集することができるため、学習に必要なデータを大幅に減らすことができる。

問題は、いかにして、クリティカルなデータを効率よく選ぶか、である。そのデータが得られたと仮定して、その事前と事後とで学習するパラメータのエントロピーが最も減るような

点を選ぶ、などの方法が提案されている。また、統計学の分野で発展してきた実験計画法とも関連が深い。

2.2. 強化学習

強化学習はゲームやロボットのための学習法としてよく知られている[3]。強化学習では、学習システムは与えられた環境の中で試行錯誤を繰り返しながら、自分の行動に対する報酬を得ることで、目的を達成する行動系列を強化・獲得してゆく。ここにおいて、システムと環境の間のインタラクションは、システムの行動とそれに対する環境からの報酬のやりとりである。

強化学習過程は、数理的にはマルコフ決定過程の同定とそこにおける最適な行動戦略の獲得として定式化することができ、その問題を逐次的に解くためのアルゴリズムとして Q 学習がよく知られている。こうした学習アルゴリズムによって、たとえば、バックギャモンでは人間のチャンピオンクラスのプレイヤーを学習によって実現することに成功している。また、最近では、ロボカップのロボットに強化学習を行わせる研究も多い。

システムにランダムな試行錯誤をさせるだけでは、複雑な環境下での目標達成をさせることはむずかしく、学習に時間がかかりすぎることが多い。そこで、強化学習の過程を加速するために、問題を少しずつ難しくしてゆく方法、問題を階層的に分解して、要素の問題から解決してゆく方法、システムにアドバイスを与える方法、などが提案されている[4,5,6,7]。また、ロボットが学習をする場合など、実際の環境で試行錯誤すると、学習が完了するまでにロボットが壊れてしまったりする。そこで、環境のシミュレータを使って学習を加速することも試みられている。

これらの中で、インタラクションという観点から興味深いのは、アドバイスを与える方法である。これは、最も簡単には、正解を教えてしまう、ということである。さらに、能動学習と

同様に、システム側が、迷っているところについてだけ正解を要求するようにすることもできるだろう。

2.3. 模倣学習

模倣学習は、知能ロボットの分野から生まれたインタラクティブな学習のスキームである。知能ロボットの分野では、抽象的なタスクを自由度の多い複雑なロボットにどのように教示するかが大きな問題の一つである。産業用ロボットでは、指定されたある軌道を精度よくこなることが重要であるが(ティーチングプレイバック)、それだけでは、作業の環境が少し変わっただけでも毎回教示をしないことになる。そこで、作業の環境やロボット自体の構成の変化に対応できるようなより汎用性の高い知識をどのようにしてロボットに与えるかが問題となる。

この問題に対して、人間がロボットの前で模範的な作業をしてみせ、ロボットがそれを模倣することによってタスクの教示を行うのが模倣学習、観察学習(Learning by Imitation; Teaching by Showing, Learning from Observation などとも言う)である。

教示は通常視覚系入力から与えられるため、視覚系における出来事を自分の運動系における出来事に変換する必要がある。また、人間とロボットの機構が異なり、タスクの初期条件なども異なるため、模範例の軌跡をそのまま座標変換するだけでは模倣はうまくゆかず、模範例をより抽象的なレベルで「理解」し、それと機能的に同等な結果が得られるように自らの運動プログラムを生成する必要がある。

実ロボットによる代表的な実装例としては、国吉による積み木作業の模倣[8]、池内と末広による組み立て作業の模倣[9]、川人による剣玉などの模倣[10]があげられる。

2.4. MAT 学習

機械学習の過程について、必要な計算量やデータ量の観点から厳密に検討する計算論的学

習理論の分野では、抽象言語の学習、特に、一定のクラスに属する文法を、正しい文章例や誤った文章例から学習することを主たるターゲットの一つとしてきたが、その中で、MAT (Minimally Adequate Teacher)学習という興味深い枠組みが提案されている。

文法学習では、能動的な質問、すなわち、ある文字列が文法に属しているか否かを問う質問(membership query)を用いた学習が主に研究されていたが、1987年に Angluin は、学習者に別のタイプの質問、すなわち、ある文法規則が正解であるか否かを問う質問(同値質問 equivalence query)を許すことで、正則文法のクラスが多項式時間で学習可能になることを示した[12]。同値質問に対して、教師は、正解である場合には正解と答えてそこで学習は終了するが、不正解である場合には、反例を示す。すなわち、正解の文法には所属し、学習者から示された文法には所属しない文の例を示す。

2.5. 進化的アプローチ

「インタラクション」と言うときに、学習システム単体と環境の間のインタラクションばかりではなく、学習システム同士のインタラクションも重要な要素である。模倣学習や対話ベースの学習では、学習者と教示者、トレーナーとの関係は固定的であったが、学習への進化的アプローチでは、そうではなくて、多数の学習者の集団を想定し、その中で学習者同士がさまざまな相互作用を繰り返す中で学習がどのように進行するかを研究対象とする。

こうしたアプローチの典型的なものは、遺伝的アルゴリズムである[13]。遺伝的アルゴリズムは最適化手法と考えられているが、最適化と学習とは裏表の関係にあり、遺伝的アルゴリズムの最適探索過程は、遺伝子の集団が、お互いに情報交換をしながら最適化の対象であるフィットネス関数の描く形を推定し、その関数によって定義される環境内での最適な振る舞いを学習してゆく過程とみなすことができる。

こうした研究に加えて、たとえば、Axerlodらの研究のように、裏切りよりは協調を重んじることが進化的に有利である、というような形で、倫理的な規範の発生を説明しようという研究、人工市場におけるディーラーの学習過程をシミュレートし、それによってバブルのような非線形な経済現象の出現を説明しようとする研究[14]、集団における言語の発生をシミュレートしようとする研究[15]など、さまざまな形で、進化過程を模擬する学習研究が行われている。

2.6. 状況論的アプローチ

学習への状況論的アプローチは、個を越えた社会的集団における学習への認知科学的なアプローチの一つであり、組織の中で知識が分散的に獲得、蓄積されてゆく様子を、集団的な作業が熟達してゆく過程の記録を取り、それを解析してゆくこと(相互行為分析)などを通じて明らかにしようとする[16]。そこでは、やはり、学習者と教示者とのダイナミックな関係や、状況における短い発話の役割り、などが解析の対象となる。

3. 対話にもとづく学習

模倣は、人間の幼児の学習においても重要な役割を果たしているが、人間の場合、それ以上に学習を加速するために重要な役割を果たしているのが言語であると考えられる。言語を介した学習というと、教科書を読んで知識を得るというような高度なものが想定されるが、そのようなもの以外にも、むしろ、タスクを実行している中でちょっとした断片的な言葉のやりとりが学習を進めるための重要なヒントになっていることが多いように感じられる。

このような感覚を機械学習の場に持ち込んでみたのが、対話にもとづく学習(対話ベースの学習)である。そこでは、ごく単純な言葉のやりとりによって、複雑なタスクや知識の学習がどのように加速されるかが興味の対象となる。

こうした方向の先駆けと言えるのは、Agreらによる研究である[17]。彼らは、学習システムにテレビゲームを学習させる過程で、同じ画面を横で見ている人間が発する断片的な言葉や画面のある地点を指示するキューなどが学習システムにどのように利用され得るかについて検討した。

一方、Torranceら[18]、および、著者ら[19]は、同様の考え方を移動ロボットの地図学習に適用した。以下、この研究について少し詳しく見てゆく。

移動ロボットにとって地図は、環境中の移動を効率よく行うために必要な重要な情報である。あらかじめ地図情報を用意してロボットに組み込むことも多いが、変動する環境に対処するため、また、ロボット自身にとって有効な地図を作成できるようにするために、ロボット自身が自ら環境を探索し、地図を作成する能力を持つことも望まれる。

移動ロボット用の地図にはさまざまな形式のものがあるが、最もよく使われているのは、

- 計量的な地図
- 位相的な地図

の二種類である。前者はいわゆる地図帳のような鳥瞰図的な地図であり、後者は、環境中のランドマークをその間の移動手段で結合したグラフとして表現される。いずれの地図も、数理的には、ロボットの現在地を隠れ状態とし、そこでの観測値を状態からの出力値とする隠れマルコフモデルとして定式化することができる。

従って、地図の学習は、隠れマルコフモデルを、状態数や状態遷移の構造も含めて学習する問題となり、一般には効率よい学習がむずかしい[20]。そのむずかしさは、ロボットが現在どの状態にいるかが不確実になるところから来る。

そこで、その学習に最も重要な情報をロボットと人間との間の対話によって補うことで、地図の学習を大幅に加速することができる。具体的には、ちょうどオフィスに新しく来たメンバ

にいろいろな場所を教えるように、環境中を人間の教師がロボットと一緒に移動しながら、ランドマークを同定するためのラベル、つまりその場所の名前を教示してゆく。

ここで用いられる対話は、ごく簡単なもので、ロボットは現在地の不確定性が高くなると「ここはどこですか?」と質問をし、それに対して教示者が場所の名前(「松井さんのところです」など)を答えるというだけのものである。

しかしながら、こうしたごく簡単な対話といえども、それを実現するためには、音声認識を始めとする多くの要素技術が必要である。また、ロボット自体が、どのような状況でどのような質問をし、返答をどのように処理するかを知っていなければならない。地図の学習のためだけにこうした大量の前提知識を仮定することはシステムの設計としては不適切であろう。

そこで、対話機能自体も、タスクと並行して学習してゆくことが望ましい。ロボットに言語を教示する試みとしては、既に、Deb Roy [21] や岩橋[22]によるものがあるが、こうした試みと対話ベースの学習とを組み合わせることは興味深い課題である。

4. 学習におけるインタラクション

さまざまなインタラクション指向の学習研究を概観してきたが、ここでは、そこにおけるインタラクションのあり方について簡単に整理してみる。

まず、インタラクションの主体については、学習システムと環境、学習システムと教示者、学習システム同士、の3種類に大きく分けてみ

る。次に、インタラクションで流れる情報の種類について、通常の種類に従って、正解情報、報酬、ヒント、その他、の4種類に分けてみる。この分類に従って、いろいろな学習スキームを当てはめてみたものが表1である。「その他」が多いことからわかるように、今後、より興味深い学習方式を得るためにも、インタラクションにおいて流れる情報をより詳しく分析する必要がある。

5. おわりに

インタラクションを通じた学習について、いろいろな分野で独立に進められている研究の概観と整理を試みた。用意されたデータを使ってオフラインで学習する静的な機械学習を超えて、社会的なインタラクションの中での学習への興味と必要性は、特に、知能ロボットや知的な対話的インターフェースのための学習の枠組みとして、今後も増大してゆくと考えられる。また、従来の、使いやすいユーザーインターフェースのための対話ではなく、システムが学習するための対話やインタラクションという切り口は、対話やインタラクション、インターフェース研究に対する新しい観点を与える可能性があるかもしれない。たとえば、普段の何気ない対話を通じて、自然にユーザの特性を学習するインターフェースなどへの応用が考えられる。本稿はまことに不完全なものであるが、そうした問題について考える際の一つのきっかけとなれば幸いである。

表1 インタラクションの主体と流れる情報

	正解	報酬	ヒント	その他
システムと環境	能動学習	強化学習		
システムと教示者	アドバイスつき能動学習	アドバイスつき能動学習	アドバイスつき能動学習	模倣学習、MAT対話ベース学習
システム同士		進化的アプローチ		進化的アプローチ 状況論アプローチ

参考文献

- [1] T. Mitchel. Machine Learning, McGraw-Hill, 1997.
- [2] 解説特集: 能動学習, 情報処理学会誌, Vol.38, No.7, 1997.
- [3] L. P. Kaelbling. Learning in Embedded Systems, The MIT Press, 1993.
- [4] J.A. Clouse and P.E. Utgoff. A teaching method for reinforcement learning. In Proceedings of the Ninth Conference on Machine Learning}, 92-101, Morgan Kaufmann, 1992.
- [5] M. Dorigo and M. Colombetti. Robot shaping: developing autonomous agents through learning, Artificial Intelligence, Vol.71, 321-370, 1994.
- [6] D. Gordon and D. Subramanian. A multistrategy learning scheme for agent knowledge acquisition. Informatica, Vol.17, 331-364, 1994.
- [7] R. Maclin and J.W. Shavlik. Creating advice-taking reinforcement learners. Machine Learning, Vol.22, 251-281, 1996
- [8] Y. Kuniyoshi, M. Inaba, and H. Inoue. Learning by watching: extracting reusable task knowledge from visual observation of human performance, IEEE Transactions on Robotics and Automation, Vol.10, 799-822, 1994.
- [9] K. Ikeuchi and T. Suehiro. Towards an assembly plan from observation, Part I: Task recognition with polyhedral objects, IEEE Transactions on Robotics and Automation, Vol.10, 1994.
- [10] M. Kawato, F. Gandolfo, H. Gomi, and Y. Wada. Teaching by showing in Kendama based on optimization principle, In Marinaro, M. and Morasso, P.G. (eds.) In Proceedings of the 1994 International Conference on Artificial Neural Networks, 601-606, 1994.
- [11] J. Demiris. Experiments towards learning by imitation. Proceedings of AAAI-94, 1437, 1994.
- [12] D. Angluin. Learning regular sets from queries and counterexamples. Information and Computation, Vol.75, 87-106, 1987.
- [13] 伊庭齊志. 「遺伝的アルゴリズムの基礎」, オーム社, 1994.
- [14] 和泉 潔. コンピュータの中の市場: 認知機構をもつエージェントからなる人工市場の構築とその評価 「認知科学」, vol6, number 1, 共立出版, 1999.
- [15] Steels, R. The talking heads experiment, <http://talking-heads.csl.sony.fr>.
- [16] 上野直樹. 「仕事の中での学習 状況論的アプローチ」, シリーズ人間の発達9, 東京大学出版会, 1999.
- [17] P. Agre and D. Chapman. Pengi: An implementation of a theory of activity. In Proceedings of the 1987 National Conference on Artificial Intelligence, 268-272, 1987.
- [18] M.C. Torrance. Natural Communication with Robots. Master thesis, MIT, 1994.
- [19] H. Asoh, Y. Motomura, T. Matsui, S. Hayamizu, and I. Hara. Acquiring probabilistic map with dialog-based learning. In Proceedings of ROBOLEARN-96, 11-18, 1996.
- [20] T. Dean, K. Basye, R. Chekaluk, S. Hyun, M. Lejter, and M. Randazza. Coping with uncertainty in a control system for navigation and exploration. In Proceedings of AAAI-90, 1010-1015, 1990.
- [21] D. Roy. Learning Words from Sights and Sounds: A Computational Model, Ph.D. thesis, MIT, 1999.
- [22] 岩橋、羽岡. 言語獲得のための視点に依存した空間概念の学習, 電子情報通信学会技術研究報告 PRMU, 2000.11.