

解説 能動学習

1. 能動学習概要

Introduction to Active Learning by Naoki ABE and Atsuyoshi NAKAMURA (C & C Research Laboratories, NEC)

安倍直樹¹ 中村篤祥¹¹ NEC C & C 研究所

1. はじめに：能動学習とは？

学習とは環境から得られる情報に基づいてその背後に潜む一般的な規則を見つけることであるが、大きく分けて以下の2つの場合がある。1つは学習者の行動が得られる情報に無関係な場合である。この場合には、学習者は環境から観察された情報のみから受動的な学習を行うしかない。ランダムに発生したデータからの学習は、このような「受動学習」の一例である。もう1つは学習者の行動が得られる情報に影響を与えられる場合である。この場合には、学習者は自らの行動を賢明に選択し、最も必要と思われるデータを得ることにより学習をより効率的に行うことができる。たとえば、実験により学習データを得られるような環境では、このような能動的な学習が可能である。このように、学習者が積極的に行動して環境から選択的に情報を得ることにより行う学習を、「能動学習 (active learning)」という⁹⁾。

学習者が取り得る行動には、大量データからの訓練データの選択、実験の計画・実行、教師への質問、環境中を探索することなど、多種多様なものがあり、「能動学習」は機械学習、情報検索、ロボット制御など多くの研究分野の重要な共通テーマとなっている。ここでは、こうしたさまざまな研究分野における能動学習の捉え方、研究活動などについて概観する。

2. 能動学習の応用

まず、能動学習の具体的なイメージを抱いてもらうために、応用について紹介する。

2.1 ロボット制御

危険な環境で有害物質を取り扱ったり、身体障害者の手助けをするなど、環境を視覚的に把握し

行動をとるロボットは、応用範囲の広い魅力的なシステムである。しかし、同じ物体でさえみる方向や光のあたり具合などにより違ってみえるだけでなく、異なる形・材質の物体が無数にあるなど、実際の応用に耐え得るためには、柔軟なシステムを構築することが必要になる。そのためには、一般的なシステムを作り、個々の応用環境に応じて自動的にカスタマイズすることが望ましい。その有力な方法の1つが、ロボットに能動的に学習させるというアプローチである。

ロボットはセンサを使って現在の状態を把握し、その状態に依存して行動をとる。そして、その行動が成功したか否かがフィードバックされ、状態から行動を決める関数としての行動規則の知識が更新 (学習) される。次の行動をとるにあたって、その時の状態におけるその関数値により行動を決めるのである。近年、強化学習の手法や決定木を用いた学習法など、さまざまな機械学習技術を適用することによりロボットの行動規則の獲得を行う研究が盛んであり、具体的な応用事例も多い^{2),11)}。

2.2 情報検索

インターネットの普及により、全世界の情報をアクセスすることが可能となったが、キーワードを指定して検索エンジンで探しても、なかなか欲しい情報を得ることができないと不満に思っている方も多いであろう。同じキーワードが入力されたとしても、人によってまたは場合によって検索すべきものは異なるし、適切なキーワードが入力されない場合もある。この問題に対処する能動学習の手法として「適合度フィードバック」(relevance feedback) というものがある。これは検索の適合度をシステムが推定し、その推定値によりランクづけされた検索結果に対して、ユーザが

その適合度の正否をフィードバックとして与えるというものである。この手法を用いれば、適合度が高いと予測されたものだけに選択的にフィードバックを要求することにより、検索者の手間をあまりかけずに個々の場合に適した検索を行うことができる。検索対象の情報か否かのラベルつきサンプルからシステムが学習し検索結果を出力するシステムと考えると、適合度フィードバックは、適合度が高いと予測されるものをシステムが選択的にサンプリングし、それに対して真のラベルを要求したとも考えられ、能動学習の一形態となっている⁶⁾。

3. 能動学習の基礎研究

上で説明した応用には、強化学習や計算論的学習理論から生み出された手法が適用されている。この章では、このような応用の種である理論研究について紹介する。

3.1 ニューラルネットワーク

ニューラルネットワークの分野においては、少ないトレーニングデータから有効な学習を行うための「選択的サンプリング」の研究が盛んである。たとえば、ベイズ決定理論の枠組みで、過去に得られた学習データから推定して、情報量が多いデータをサンプリングするための手法の研究など⁷⁾がある。また、領域の学習において、過去に得られたトレーニングデータから推定して境界と思われる領域のデータをサンプリングする手法に関する研究も盛んである⁴⁾。この分野の研究の詳細に関しては、本特集の解説2と解説3を参照されたい。

3.2 強化学習

最近、機械学習の分野で注目されている「強化学習」は、能動学習の典型的な形態である⁶⁾。強化学習では、学習者は自分のいる状態からの関数として行動を決定しその結果として別の状態に移る。その際、各行動の後に行動の善し悪しに応じて「強化信号」として報酬が得られ、それに依りて自分の行動規則を学習しなおす。学習者の目標は、将来まで考慮して報酬をできるだけ多くもらうように行動することである。強化学習の手法としては、「Q学習法」などいくつかの手法が知られている。Q学習法は、各状態の価値をその状態から期待される累積報酬(Q-value)として定量

化し、このQ-valueを逐次的に更新することにより学習していくもので、漸近的に最適な行動規則に収束することなどが保証されている⁶⁾。強化学習の枠組みはロボットの学習を定式化したものになっており、これらの学習手法のロボット制御分野への応用研究が盛んである。強化学習のロボット制御への応用については、本特集の解説5を参照してほしい。

3.3 計算論的学習理論

計算論的学習理論の分野では、選択的サンプリングのモデルとして Angluin¹⁾によって定式化された「所属性質問による概念/言語学習」の枠組み内での研究が盛んである。とくに、DNF (Disjunctive Normal Form: 積和形式)を始めとするブール関数の表現形、そして DFA (Deterministic Finite-state Automaton: 決定性有限オートマトン)などの形式言語の表現形の学習について多くの研究がなされている。ここで、「所属性質問」とは、学習者が選んだ任意の点 x における学習対象関数の値 $f(x)$ を得ることができる質問である。DFAを含むいくつかの重要な表現形のクラスについて、ランダムサンプルのみからは効率的な(多項式時間内で)学習が可能でないが、所属性質問を許せば効率的な学習が可能になることが判明している。

このほか、この分野で研究されているもう1つの学習モデルである「逐次型学習モデル」の枠組みを用いて、前記の強化学習方式の理論的解析も行われている¹²⁾。ここでは、先に述べたような漸近的な評価ではなく、知識表現としての最適な行動規則との誤差がどの程度になるかが、有限試行回数において評価されている。また、行動を通して得られた利潤の最大化というさらに現実的な視点から眺め、ある単純化された問題について解析を行ったのが文献3)の「最適選択の逐次型学習」の研究である。ここでは、累積利潤の最大化の系として、行動規則獲得のために必要な行動と利潤の最大化のために望ましい行動との間のトレードオフの解決に関する理論的解析がされている。

4. なぜ能動学習か?

4.1 サンプリングコストの削減

能動学習の最も明白なメリットは、サンプリン

グコストの削減である。多くの応用において、学習データの収集はコストが高く、高水準の学習精度を達成するための大きなネックとなっている。とくに、ラベルつきデータは人手が必要なためコストが高い。テキスト分類を例にとると、分類クラスラベル付きの文書データなどには人手による評価が不可欠である。実際、この分野では“Text is cheap, but information is costly”⁷⁾などといわれている。この場合、ランダムに選んだ文書に対して分類してもらうのでは、コストパフォーマンスが悪い。この問題に対しては、現在の仮説での分類クラスラベルの推定が最も曖昧な文書についての実際のラベルを教えてもらい、逐次的に仮説を改良するような能動学習法 (uncertainty sampling) を用いることにより、人が分類しなければならぬデータを大幅 (最高 500 分の 1 程度) に削減できたという結果が報告されている⁷⁾。

4.2 計算コストの削減

学習を行うのに十分なデータがある場合でも、そのデータから精度のよい規則を導くことが計算コスト的に困難である場合がある。たとえば、正則言語の表現形として知られている DFA について、正例と負例の文字列をいくつか与えられた場合、それに矛盾しない小さな DFA を求めることは計算コスト的に困難であることが示されている¹⁰⁾。一般に、ランダムサンプルから確率的に近似的に正しい学習 (いわゆる PAC 学習¹³⁾) が可能なことと、与えられたデータに対して無矛盾で小さな仮説を求められることはほぼ同値であるので¹⁰⁾、ランダムサンプルからの DFA の学習は計算コスト的に困難であることが分かる。

これに対して、「この文字列は正例か負例か？」という質問 (前出の所属性質問) に答えてくれる教師が存在する場合には、少ない計算コストで与えられたランダムサンプルに無矛盾な最小の DFA を学習するアルゴリズムが存在する¹¹⁾。この例が示すように、能動学習はサンプリングコストの削減にとどまらず、受動学習が抱える計算コストの問題を解決する有効な手段となりうるのである。

4.3 総合的な利潤の最大化

与えられたデータの中からその背後に潜む規則を発見することを目的とする「受動学習」は、音

声認識、パターン認識などといった個々の知的機能の獲得のモデル化として有効である。しかし、ロボットのように環境中で「行動」をとることにより環境について「学習」をして、また「行動」をとることにより「報酬」を得るような「知的行動者」の統合的なモデル化はその範疇にない。ロボットのモデル化としての能動学習の枠組み (強化学習) においては、単なる関数や概念の学習のモデル化ではなく、行動規則の獲得のための行動と獲得された知識の利用による報酬の最大化の両方が考慮されている。このような能動学習の枠組みを用いることにより、サンプリングコスト削減や計算コスト削減の技術も要素的技術として活かしつつ、総合的な利潤の最大化のための独特の技術が開発されることが期待できる。

5. おわりに

近年「能動学習」と呼ばれ注目されている技術に関する研究活動について概観した。さまざまな分野のさまざまな研究が、能動学習という視点から統一的に眺めることができる。今後、重要な共通テーマの統合的な研究や他分野の成果の応用など、分野をまたがって技術者・研究者が交流し成果を共有することにより、有益な技術が広く応用されることを望みたい。

参考文献

- 1) Angluin, D.: Learning Regular Sets from Queries and Counterexamples, *Information and Computation* 75, pp. 87-106 (1987).
- 2) Asada, M., Noda, S., Tawaratsumida, S. and Hosoda, K.: Purposive Behavior Acquisition for a Real Robot by Vision-Based Reinforcement Learning, *Machine Learning* 23, pp. 279-303 (1996).
- 3) Abe, N. and Takeuchi, J.: The 'Lob-Pass' Problem and an On-Line Learning Model of Rational Choice, *Proceedings of the 6th Annual Workshop on Computational Learning Theory*, pp. 422-428 (1993).
- 4) Baum, E. B.: Neural Net Algorithms That Learn in Polynomial Time from Examples and Queries, *IEEE Transactions on Neural Networks* 2 (1), pp. 5-19 (1991).
- 5) Chon, D. and Lewis, D. D.: (The 1995 Fall Symposia Series) Active Learning Symposium, *AI Magazine*, p. 83 (spring 1996).
- 6) Kaelbling, L. P., edit.: A Special Issue on Reinforcement Learning, *Machine Learning*, 22 (1/2/3) (1996).

- 7) Lewis, D. D. and Gale, W. A. : A Sequential Algorithm for Training Text Classifiers, In Proceedings of the 17th Annual International ACM-SIGIR Conference on Research and Development in Information Retrieval, pp. 3-12 (1994).
- 8) Lewis, D. D. : Active by Accident : Relevance Feedback in Information Retrieval, Unpublished Working Note of the 1995 AAAI Fall Symposium on Active Learning (1995).
- 9) Paass, G. and Kindermann, J. : Bayesian Query Construction for Neural NetWork Models, Advances in Neural Information Processing Systems 7, pp. 443-450 (1995).
- 10) Pitt, L. and Warmuth, M. K. : The Minimum Consistent DFA Problem Cannot be Approximated within any Polynomial, J. A. C. M., 40 (1), pp. 95-142 (1993).
- 11) Salganicoff, M., Ungar, L. H. and Bajcsy, R. : Active Learning for Vision-Based Robot Grasping, Machine Learning 23, pp. 251-278 (1996).
- 12) Schapire, R. E. and Warmuth, M. K. : On the Worst-Case Analysis of Temporal-Difference Learning Algorithms, Machine Learning, pp. 95-121 (1996).
- 13) Valiant, L. G. : A Theory of the Learnable, Communications of the A. C. M., 27, pp. 1134-1142 (1984).

(平成9年5月1日受付)



安倍 直樹 (正会員)

1960年生, 1984年マサチューセツ工科大学情報科学科学士/修士課程修了, 同年IBMワトソン研究所にて研究員, 1989年ペンシルバニア大学情報科学科博士課程修了, 同年カリフォルニア大学サンタクルーズ校にて研究員, 1990年NEC入社, 現在C & C研究所研究専門課長, 計算論的学習理論およびその応用研究に従事, Ph.D.
e-mail : abe@sbl.cl.nec.co.jp



中村 篤祥 (正会員)

1963年生, 1986年東京工業大学理学部情報科学科卒業, 1988年同大学院理工学研究科修士課程修了, 同年NEC入社, 現在, 同社C & C研究所に勤務, 計算論的学習理論の研究に従事, e-mail : atsu@sbl.cl.nec.co.jp