

# 統計的学習はロボットを救うか？

Can Statistical Learning Help Robots?

麻生英樹

Hideki ASOH

独立行政法人 産業技術総合研究所

National Institute of Advanced Industrial Science and Technology, AIST

It is considered that learning is indispensable for making robot intelligent. In this talk, I will give an overview of some researches on applying statistical learning to intelligent robots, and try to clarify central issues, possible contributions, and ways to a breakthrough.

## 1. はじめに

ロボットは、WWW や個人情報端末などと並んで、人工知能研究の重要なプラットフォームの1つである。すなわち、多くの人工知能研究の成果は、実環境で働くロボットという場において、その真価を試される。様々なペットロボットや二足歩行ロボットが開発されるなど、ハードウェア的な進歩が著しい現在、そうしたロボットをどれだけ「知的に」できるかということは、工学的なあるいは産業的な意義が未知数であるとしても、人工知能の研究者にとって大変興味深い挑戦である。

ロボットを知的にするにあたって、経験やデータからの学習が重要な役割を果たすことは明らかであろう。この世界は、曖昧さ、不確実性に満ちている。すべてを自明なこととして予測するラプラスの悪魔は、熱力学的に不可能であることが証明されているし、カオス的な力学系が原理的に予測不可能であることも明らかになっているが、たとえそうでなくても、情報処理能力の現実的な限界があるのは明らかである。そして、にもかかわらず、というよりも、だからこそ、我々はできる限り、未来を予測しようとし、そのために、過去の経験から学ぼうとする。また、実際的にも、複雑な環境で働くために、あらゆる場面や環境を想定してあらかじめプログラムを書いておくことは難しい。そこで、人工知能研究の初期から、生物に習って、学習プログラム、すなわち、自らの経験によって、プログラムを調整したり合成したりするためのメタプログラム、を組み込んでおく戦略が探求されてきた。

このような機械学習の研究の中でも、確率・統計的な手法を利用する統計的学習の研究は、大量の学習データに基いた対象の認識、分類、識別の問題を中心に大きな成功を納めてきている [Hastie 01, Vapnik 95]。本発表では、統計的学習が知能ロボット研究にどのように寄与するかについて考える。

以下では、まず、これまでの、統計的学習をロボットに適用した研究の中から、代表的と著者が考える研究例を紹介し、それらに共通する課題が何であるのかを明らかにするとともに、統計的学習がその課題にどのように寄与できるのか、また、現在の状態を越えてゆける可能性があるのか、について考察する。

## 2. 学習するロボットたち

「学習するロボット」という言葉を始めて目にしたのは鉄腕アトムの中だったか鉄人 28 号の中だったか記憶が朧なのだが、「このロボットは経験から学んで賢くなってゆく」という

ような台詞と、しかし、ロボットに何かを教えるのは難しく、ロボットに良いこととして教えたことが状況が変わって悪いことになってしまう、というストーリーだけが強く印象に残っている。

それから数十年が過ぎた現在、多くの「学習するロボット」が研究され、実際に作られたロボットが稼働している。小嶋らのインファノイドプロジェクト [小嶋 01] や浅田らの認知発達ロボティクス [浅田 03] などに代表されるように、ロボットを使って人間や動物の認知発達過程の研究を行うことも盛んになりつつある。

ここでは、そうした研究の中から、行動の強化学習、環境地図の学習、見まね（模倣）学習、環境相互作用ダイナミクスの学習、言語使用とコミュニケーションの学習、の4つを取り上げ、それぞれにおいて学習がどのように働いているかを概観する。なお、研究の選択は著者の限られた知見に基づくもので、ここで示すもの以外にも多くの興味深い研究があるのは言うまでもない。

### 2.1 行動の強化学習

強化学習 [Sutton 98] はロボットの学習において最もポピュラーな手法の1つで、迷路抜けや障害物回避から対話制御の最適化 [Singh 99, Roy 00] などまで、多くの課題に適用されて研究されている。そのアルゴリズムの起源は動物の行動学習の研究に遡り、最適制御問題、不確実環境化での最適行動選択問題としての定式化を得ることによって数理的に洗練された。また、最近では、実際の脳の大脳基底核においても強化学習のために必要とされる報酬信号や報酬予測誤差信号が表現されているといわれており、様々な仮説の提案と検証が進められている（たとえば [Doya 00] を参照）。

工学的な研究においては、学習に時間がかかることが主たる問題点とされ、学習を効率化するための工夫が多く提案されてきた。そのためのポイントの1つは、問題を表現するための内部状態表現やセンサ情報表現の設計である。ロボットのハードウェアが決まれば、センサとして何が使えるかは所与となるが、それでも、特に視覚センサのような情報量の多いセンサの場合には、それが捉える情報から何を重要な情報として抽出してロボットの内部状態表現を構成するかには大きな自由度がある。これは、パターン認識における特徴抽出と同じ問題であり、多くの場合に、タスクに適した情報を抽出できているか否かが学習の成否を左右する。

現時点では、このような内部状態空間の階層的な作り込みに加えて、学習課題の階層化、学習環境の上手な設定、自由度を漸増させる手法、などの工夫を組み合わせることにより、かなり効率の良い行動学習が可能になってきているが、それと並

連絡先: 麻生英樹, 産業技術総合研究所情報処理研究部門, 〒305-8568 つくば市梅園 1-1-1 中央第2, h.asoh@aist.go.jp

行して、状態空間の構造を学習させる試みも行われている。たとえば浅田らは、その研究の初期より、タスクと環境とロボットシステムの3つの間の相互作用から適切な内部表象が創発する過程に注目して研究を行っているが、この問題に対しても、強化学習のための内部状態表現を経験を通じて獲得する手法を検討し、単純に環境をセンサ情報の空間での近さだけで分類するのではなく、同じ種類の行動をすることによって同じ結果を得られる領域をひとつのまとまりとしてゆくことで、センサ情報の空間と行動コマンドの空間とを、同時並行的に、タスク依存な形でセグメンテーションできることを示している[浅田 97]。また、上野らもほぼ同時期に同様の研究を行っている[上野 00]。

## 2.2 環境地図の学習

強化学習による行動学習と並んで、移動ロボットのための環境地図の学習も長年にわたって研究されてきたテーマの1つである[Basye 89, Thrun 98a]。移動ロボットの自己位置の同定には地図が必要であり、精度のよい地図の作成には自己位置の同定が必要であるが、この2つの問題を同時に解決する問題は、SLAM(Simultaneous Localization and Mapping)と呼ばれて、特に盛んに研究されてきた[Thrun 98b]。そこでは、環境同定型の強化学習と同様に、ロボットと環境とのインタラクションの過程を隠れマルコフ過程としてモデル化することが行われ、地図の学習は、隠れマルコフモデルの推定問題に帰着される、この意味では、強化学習から特定の目的達成と連合した報酬を除いた部分問題と考えることもできる。また、逆に、新規な場所の発見や地図への追加を報酬として未知領域を探索させる研究もある。現時点では、レーザーレンジファインダやステレオカメラなど、情報量の多いセンサを用いて、ほぼリアルタイムの地図作成ができるようになっており、戦闘地域の偵察や惑星探査などにおいても実用的なレベルにまで達している。

ここでも、学習の鍵となるのは、センサ情報の表現、ロボットの行動情報の表現、およびそれらの関係としての環境地図の内部表現の設計である。地図の表現としては、必要な次元のメッシュグリッドを用意し、各グリッドに情報(たとえば、そこに物がある、など)を書き込んでゆくことによって環境の物理構造をほぼそのまま表現するものがある[Elfes 92]。この方法は、物理的空間の位相を保存することができるため、メトリカルな地図などと呼ばれている。メッシュを細かくし、各メッシュに記述できる属性を増やしてゆけば、非常に詳細な地図を表現することができるが、その分、情報量が多く、経路計画などに利用するときには、そこからさらに適切な情報を抽出する必要がある。これに対して、重要な場所(ランドマーク)だけをノードとし、その間の接続関係や移動関係をアークによって表現するグラフ表現も用いられ、メトリカルな地図に対して、トポロジカルな地図などと呼ばれている[Kuipers 88]。さらに、これらの中間のような地図や、場所の表現を抽象化することで複数の表現を融合したような地図も考えられている。センサ情報の表現の獲得についてはあまり多くの研究は無いが、たとえばThrunらは、自己位置同定の精度を評価尺度とした学習を行わせて、センサ情報から重要な情報を特徴抽出させている[Thrun 98c]。

## 2.3 見まね学習

ロボットによる模倣学習、あるいはteaching by showingの研究は、国吉[Kuniyoshi 94]や池内、末広らの研究[Ikeuchi 94]などをその嚆矢としているが、赤ん坊の認知発達過程における模倣の重要性の認識や、脳内におけるいわゆるミラー・ニュー

ロン(自分がある特定の行為をしている時だけでなく、他者が同じ行為をしているのを観察したときにも発火する神経細胞)の発見などもあいまって、現在では多くのロボット研究者がその重要性を指摘して研究を行っている。

その中でも、川人らのヒューマノイドロボットDB-Dynamic Brainは、それが学習した芸の高度さと数の多さによって最もよく知られているものの1つである[DB 99, Schaal 99]。その一部は2001年度の本学会全国大会特別講演の中でも紹介されたが、ジャグリング、パドリング、ポールバランス、剣玉、エア・ホッケー、踊り等を見まね学習する姿は大変印象的であった。

こうした見まね学習を可能にするためには、1)自分の身体や道具の内部モデルの内部表現、2)視覚等のセンサを通じて獲得される相手の身体および道具の状態の表現、3)2つの表現の関係についての知識、が必要となるが、多くのシステムにおいては、これらの仕組みを設計者が作り込み、それを利用してロボットに教師である人間の行為の模倣をさせている。

これに対して、こうした内部表現やそれらの関係自体をも経験から獲得させようという研究もある。たとえば、国吉らは、自分の手の動きを観察することを通じて、視覚からの情報と、自分の身体の運動の関係を学習させることを提案している。また、浅田らは、共同注視の学習において、1つレベルの低い単独注視の成功をキーとする学習を行うことで、教師無しでも養育者の顔画像から適切な運動コマンドを生成できるようになることを示している[浅田 03]。

また、現在のシステムの多くでは、模倣対象の行為はあらかじめ切り出されていることが多いが、実際に現実の状況の中で模倣学習をするためには、一連の感覚運動データを、模倣しやすいように適切に分節する必要もある。この問題については、次節でも触れる。

## 2.4 環境との複雑な相互作用ダイナミクスの分節と学習

環境地図の学習や見まね学習では、環境や道具の特性は比較的固定的、静的なものと考えられていたが、環境自体が複雑なダイナミクスを持っているような場合や、複数の道具を切り替えて使用するような場合、には、一連の感覚入力や運動出力の流れを適切に分節しながら、要素となるダイナミクスの学習を進めてゆく必要がある。

谷らは、移動ロボットの環境地図学習から始まる一連の研究において、回帰結合型のニューラルネットワークによって複数の非線形ダイナミクスを学習させる方法について検討している[Tani 95, Tani 99]。川人らも、順モデル(予測器)と逆モデル(制御器)とを組み合わせたペアをモジュールとして多数用意し、順モデルによる予測の当たり具合によって切り替えて使うMOSAICアーキテクチャを提案し、それを階層的に組み合わせて用いることによって、様々な道具使用の学習からコミュニケーションの学習まで幅広く応用することを提案している[Haruno 01]。

こうした研究は、予測の当たり具合という教師無しでも求められる評価量を鍵として、複雑なダイナミクスを、ダイナミクスとしての類似性という観点から分節、分類して学習しようとするものであり、現時点ではまだまだ原初的な段階であるが、今後の進展が期待される。

## 2.5 言語使用の学習

人間の幼児の言語獲得過程については、多くの認知科学的な研究が行われているが[Saffran 96, 今井 00]、ロボットに言語使用を学習させる研究も始まっている。たとえば、Royらは、画像と音声の連合を利用して、ロボットに語彙や文法を獲

得させることを試み、「これはリンゴ」というような音声の提示と、対象物の視覚的な提示とを連合させることで、ロボットに語彙や簡単な文法の学習を行わせ、「リンゴを見て」というような命令に対応できるようにした [Roy 99]。画像と音声の関係を学習するために、そこでは、相互情報量を最大化する基準を導入し、それぞれのモダリティの情報をクラスタリングしている。実際の母子のインタラクションから収集したデータを用いてロボットに学習させた結果も示している。

岩橋らは、同様の課題に取り組み、音韻レベルのクラスタリングの獲得を試みるとともに、物体の移動軌跡も含めて学習させることで、動詞の学習を可能にしたり、相手の意図を推測することによって省略への対応を学習させたりしている [岩橋 03a, 岩橋 03b]。

言語の使用、対話、より広く非言語的な情報伝達まで含めたコミュニケーションの過程を、他者を含む環境との相互作用ダイナミクスを利用して、他者および自分と他者との関係を理解・制御してゆく問題として捉え、学習させることも研究されている [Paek 00] さらに、ロボットを用いて言語の創発過程を探求している研究もある [Steels 98]。

### 3. 考察：統計的学習はロボットを救うか？

「統計的学習はロボットを救うか？」という問いへの答えを最初に言ってしまうと「それだけでは救えない。しかし、役には立つだろう」という、ごくあたりまえのものになる。そもそも、ある手法やアプローチが、それだけでロボットや人工知能研究を救う、などということは、あまり考えないほうがよい。しかし、そうだからこそ、データを集めて学習させさせればなんとかなる、というような過剰な期待や、学習できることなどほんの少しだ、というような過小な評価を排して、上のあたりまえの答えが含意することをよく理解することは重要だろう。

上に示したいくつかの例に共通している本質的な問題の1つは、観測された状態や内部状態や状況、タスクのゴールなどを分節、クラスタリング、粗視化、階層化し、効率よく利用できる形で表現することである。それらは、生物の進化の過程においては、長い時間をかけて獲得されてきたものであり、オントロジーの構築などとも関連する、人工知能の基本問題の1つと言える。

進化の過程に見合うような十分な時間をかければ、すべてを学習させることが可能かもしれないが、それは明らかに現実的ではないだろう。学習に関与する十分な量のデータが収集できれば、それを高速に体験させることが考えられるが、現実には、実際の世界で起こるインタラクションのうち学習に関与する部分だけを切り離すことは難しいため、世界全体をシミュレートすることが必要になってしまう。

従って、何を生得的に組み込み、何を学習させるか、が設計上の問題となるのだが、この点についての見通しはほとんど得られていないと言えるだろう。個々の研究では、特定の問題(群)が切り取られ、学習を成功させるに必要な設定や技巧が探求されているが、感覚レベルから言語レベルに至るまでのさまざまな学習を総合的に実現しているシステムはまだ存在していないと思われる。ただ、MOSAIC を中心とした構想に見られるように、そうした方向への研究も始まりつつある。

上の情報表現の問題は、より広い観点から、学習のための情報をどのようにシステムに供給するか、という問題の一種と捉えることもできる。適切な情報表現や学習のためのモデル、学習環境、タスク群などを設計することは、生物の認知科学的な研究、脳科学的な研究、あるいは、過去の機械学習研究から得

られた知見を学習システムに供給することにあたる。統計的学習では、データが主たる情報源となっているが、実際的には、上記のようなデータ以外から得られる情報の果たしている役割は非常に大きい。

そこで、より積極的に、現在の統計的学習におけるデータのあり方、使い方を見直したり、データとデータ以外からの知識とを融合してゆく方法を模索したりすることも行われている。たとえば、統計的学習が最も高度に利用されている分野の1つである音声認識技術においては、汎用的な学習のための大量のデータと、個別話者に対する少量のデータを組み合わせることで、個別話者に対する性能を大幅に向上させることに成功している [Leggetter 95]。また、音声認識における言語モデルの学習において、学習用の言語データと、固定的な言い回しの部分に対する文法知識とを融合する方法の提案もある [Akiba 03]。ロボットの強化学習においても、教師からのアドバイスを利用する手法が多数提案されている [Maclin 96]。我々も、地図学習を例題として、統計的学習をロボットと人間との間の社会的なインタラクションの中に置き、センサから得られる統計的なデータに加えて、人間との対話から得られる状況依存的な少量の情報を利用することを提案した [Asoh 97, 麻生 03]。

### 4. おわりに

いくつかの具体的な研究事例を通じて、ロボットにおける統計的学習のあり方と可能性について考察した。確率統計的な手法は、矛盾を許容し、頑健なシステムを構築することを可能にするとともに、その振る舞いをデータから学習させることを可能にする。中核となる情報表現の設計や学習の枠組み自体が学習によって解決されるということはいえそうにないが、そうした設計を補完する手段として、確率統計的手法、統計的学習は、知的なシステムの実現のために必要不可欠なものであると考えている。

今後、学習課題が複雑になるに従い、学習のための情報全体をどのような形でシステムに供給すべきか、という観点から、学習の枠組みを再検討し、総合的に設計することが重要になってゆくと考えている。その中において、何を生得的に組み込み、何を学習させるか、という問題が具体的に解決されてゆくことを期待したい。

謝辞：本研究の一部は科研費 14208033 による。

### 参考文献

- [Akiba 03] Akiba, T., Itou, K., and Fujii, A.: Adaptive language models for frequent fixed phrases by emphasizing N-gram subsets, *Proceedings of 8th European Conference on Speech Communication and Technology* (2003).
- [浅田 97] 浅田 稔, 野田 彰一, 細田 耕: ロボットの行動獲得のための状態空間の自律的構成, *日本ロボット学会誌*, vol.15, pp.886-892 (1997).
- [浅田 03] 浅田 稔: 認知発達ロボティクスにおける学習, *人工知能学会誌*, vol.18, pp.550-558 (2003).
- [Asoh 97] Asoh, H., Hayamizu, S., Hara, I., Motomura, Y., Akaho, S., and Matsui, T.: Socially embedded learning of the office-conversant robot Jijo-2, *Proceedings of the 15th International Joint Conference on Artificial Intelligence (IJCAI-97)*, pp.880-885 (1997).
- [麻生 03] 麻生英樹: 機械の学習と人間の学習-統計的学習の立場から- *人工知能学会誌*, vol.18, pp.526-530 (2003).

- [Basye 89] Basye, K., Dean, T., Vitter, J. S.: Coping with uncertainty in map learning, *Proceedings of the 11th International Joint Conference on Artificial Intelligence*, pp.347–352 (1989).
- [DB 99] <http://www.kawato.jst.go.jp/DB/home.html> (1999).
- [Doya 00] Doya, K.: Complementary roles of basal ganglia and cerebellum in learning and motor control, *Current Opinion in Neurobiology*, vol.10, pp.732–739 (2000).
- [Elman 96] Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., and Plunkett, K.: *Rethinking Innateness, A Connectionist Perspective on Development*, MIT Press (1996) (乾, 今井, 山下 (訳): 認知発達と生得性, 心はどこから来るのか, 共立出版 (1998)).
- [Elfes 92] A. Elfes: Multi-source spatial data fusion using Bayesian reasoning, in M.A.Arbib and R.C.Gonzales eds., *Data Fusion in Robotics and Machine Intelligence*, 137-163 Academic Press (1992).
- [Haruno 01] Haruno, M., Wolpert, D.M., Kawato, M.: Mosaic model for sensorimotor learning and control, *Neural Computing*, vol.13, pp.2201–2222 (2001).
- [Hastie 01] Hastie, T., Tibshirani, R., and Friedman, J.: *The Elements of Statistical Learning — Data Mining, Inference, and Prediction*, Springer-Verlag (2001).
- [Ikeuchi 94] Ikeuchi, K. and Suehiro, T.: Toward an assembly plan from observation, Part I: Task recognition with polyhedral objects, *IEEE Trans. Robotics and Automation*, vol.10, pp.356–385 (1994).
- [今井 00] 今井むつみ (編著): 心の生得性 言語・概念獲得に生得的制約は必要か, 共立出版 (2000).
- [岩橋 03a] 岩橋直人: ロボットによる言語獲得—言語処理の新しいパラダイムを目指して—, *人工知能学会誌*, vol.18, pp.49–58 (2003).
- [岩橋 03b] 岩橋直人: 言語コミュニケーションのための機械学習, *人工知能学会誌*, vol.18, pp.522–525 (2003).
- [Kawato 94] Kawato, M., Gandolfo, F., Gomi, H., and Wada, Y.: Teaching by showing in Kendama based on optimization principle, *Proceedings of the 1994 International Conference on Artificial Neural Networks*, pp.601–606 (1994).
- [小嶋 01] 小嶋秀樹, 高田 明: 社会的相互行為への発達のアプローチ—社会の中で発達するロボットの可能性, *人工知能学会誌*, vol.16, pp.812–818 (2001).
- [Kuipers 88] G.J. Kuipers and Y.-T. Byun: A robust, qualitative method for robot spatial reasoning. In *Proceedings of the AAAI-88*, 774-779 (1988).
- [Kuniyoshi 94] Kuniyoshi, Y., Inaba, M., and Inoue, H.: Learning by watching: extracting reusable task knowledge from visual observation of human performance, *IEEE Transactions on Robotics and Automation*, Vol.10, pp.799–822 (1994).
- [Leggetter 95] Leggetter, C. J. and Woodland, P. C.: Maximum likelihood linear regression for speaker adaptation of continuous density hidden Markov models, *Computer Speech and Language*, vol.9, pp.171–185 (1995).
- [Levin 00] Levin, E., Pieraccini, R., and Eckert, W.: A stochastic model of human-machine interaction for learning dialog strategies, *IEEE Transactions on Speech and Audio Processing*, Vol.8(1), pp.11-23 (2000).
- [Maclin 96] R. Maclin and J.W. Shavlik. Creating advice-taking reinforcement learners. *Machine Learning*, vol.22, pp.251-281, (1996).
- [Paek 00] Paek, T. and Horvitz, E.: Conversation as action under uncertainty, *Proceedings of the 16th Conference on Uncertainty in Artificial Intelligence (UAI-2000)* (2000).
- [Roy 99] Roy, D.: *Learning Words from Sights and Sounds: A Computational Model*, Ph.D. thesis, MIT (1999).
- [Roy 00] Roy, N., Pineau, J., and Thrun, S.: Spoken dialog management for robots, *Proceedings of the 38th Annual Meeting of the Association for Computational Linguistics (ACL2000)* (2000).
- [Saffran 96] Saffran, J. R., Aslin, R. N., and Newport, E. L.: Statistical learning by 8-month-old infants, *Science*, Vol.274, pp.1926–1928 (1996).
- [Schaal 99] Schaal, S.: Is imitation learning the route to humanoid robots?, *Trends in Cognitive Science*, vol.3, pp.233–242 (1999).
- [Singh 99] Singh, S., Kearns, M., Litman, D., and Walker, M.: Reinforcement learning for spoken dialog systems, *NIPS99* (1999).
- [Steels 98] Steels, L.: The origin of syntax in visually grounded robotic agents, *Artificial Intelligence*, vol.103, pp.1–24 (1998).
- [Sutton 98] Sutton, R. S. and Barto, A. G.: *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA (1998) (三上貞芳, 皆川雅章 訳. 強化学習. 森北出版 (2000)).
- [Tani 95] Tani, J.: Self-organization of symbolic processes through interactions with the physical world, *Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI-95)* (1995). pp. 241-249 (1995)
- [Tani 99] Tani, J. and Nolfi, S.: Learning to perceive the world as articulated: an approach for hierarchical learning in sensory-motor systems, *Neural Networks*, vol.12, pp.1131–1141 (1999).
- [Thrun 98a] Thrun, S.: Learning metric-topological maps for indoor mobile robot navigation, *Artificial Intelligence*, Vol.99(1), pp.21–71 (1998).
- [Thrun 98b] Thrun, S, Fox, D. and Burgard, W.: A probabilistic approach to concurrent mapping and localization for mobile robots, *Autonomous Robots*, vol.5, pp.253–271 (1998).
- [Thrun 98c] Thrun, S.: Bayesian landmark learning for mobile robot localization, *Machine Learning*, vol.33, pp.41–76 (1998).
- [上野 00] 上野敦志, 中須賀真一, 堀 浩一: 自律エージェントのための状況認識と行動規則の同時学習, *人工知能学会誌*, vol.15, pp.297–308 (2000).
- [Vapnik 95] Vapnik, V.: *The Nature of Statistical Learning Theory*, Springer-Verlag (1995).