Toward a Theory of Embodied Statistical Learning

Daniel Burfoot[1], Max Lungarella[2], and Yasuo Kuniyoshi[1]

[1] The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan
{burfoot, kuniyoshi}@isi.imi.i.u-tokyo.ac.jp

[2] Artificial Intelligence Laboratory, Andreasstrasse 15, 8050 Zurich, Switzerland
lunga@ifi.uzh.ch

Abstract. The purpose of this paper is to outline a new formulation of statistical learning that will be more useful and relevant to the field of robotics. The primary motivation for this new perspective is the mismatch between the form of data assumed by current statistical learning algorithms, and the form of data that is actually generated by robotic systems. Specifically, robotic systems generate a vast unlabeled data stream, while most current algorithms are designed to handle limited numbers of discrete, labeled, independent and identically distributed samples. We argue that there is only one meaningful unsupervised learning process that can be applied to a vast data stream: adaptive compression. The compression rate can be used to compare different techniques, and statistical models obtained through adaptive compression should also be useful for other tasks.

1 Introduction

One striking characteristic of human competence is that it requires many years of learning to develop. Learning can be regarded as a form of *statistical adaptation* in which the brain adjusts to data flowing into it from the senses. Recently, researchers in the field of statistical learning have made important progress in understanding the nature of learning and the conditions under which learning can occur. This understanding supports the definition of several powerful learning algorithms [1, 2].

The field of embodied artificial intelligence is also deeply concerned with the issue of adaptation, and has recently made several important conceptual advances [3]. One such advance is the realization that in many cases good performance can be achieved without advanced information processing, by relying on techniques such as reactivity, self-organization, and exploitation of body dynamics [4, 5]. Another achievement is the identification of a set of design principles to guide the construction of robotic systems [3].

Unfortunately, there is not much communication between these two disciplines. In particular, it is difficult for roboticists to apply the strong results of statistical learning theory to embodied agents research. This difficulty is caused by a mismatch between the form of data assumed by current statistical learning algorithms and the form of data available to embodied agents. The purpose of this paper is to argue for a new formulation of statistical learning that can be applied to the vast unlabeled data stream generated by robotic systems. Furthermore, we argue that given this type of input, there is only one meaningful learning process that can be applied: adaptive compression. An important advantage of the view of learning as compression is that it provides a rigorous and highly practical research methodology within which to proceed. We refer to the hybrid field resulting from a combination of ideas from statistical learning and embodied artificial intelligence as “Embodied Statistical Learning” (EStL).

2 Background

2.1 Statistical Learning: the Current Formulation

The goal of the field of statistical learning is to discover algorithms which build statistical models from data. This field has developed an impressive mathematical theory [2, 6, 7] and has demonstrated strong results on various applications, such as face detection, handwritten digit recognition, and machine translation. The basic problem statement of statistical learning, in its current form, is given in the first sentence of the first chapter of the great work by Vapnik [6]:

> In this book we consider the learning problem as a problem of finding a desired dependence using a *limited* number of observations.

Several important ideas are contained in this statement. First, a critical aspect of this type of learning is the limitations on the amount of available data. Second, the goal is to find dependencies - for example, finding a rule that can assign a label to an image (e.g. “face” or “no face”). Third, the data is assumed to be partitioned into a number of distinct “observations”. Because of this assumption of partitionability, it is then natural to assume that the samples are independent and identically distributed (IID). A critical piece of the VC theory is a set of probabilistic bounds on the difference between the real and empirical performance of a model class, in terms of the complexity of the class and the number of observations. These bounds are obtained using the assumption of IID samples, and their purpose is to describe when it is possible to generalize from limited data [6]. Thus, the above assumptions are essential to the theory. Other formulations of statistical learning are mostly similar to Vapnik's; we refer to these collectively as CStL.

2.2 Embodied Artificial Intelligence

The subfield known as embodied artificial intelligence originated with the work of Brooks [8, 9]. Writing in the early 1990s, Brooks was reacting to what he saw as an unhealthy overemphasis on the *physical symbol system hypothesis*, which was a major influence on AI at the time. In this view, the key role of intelligence was to use formal logic and symbolic manipulation to construct plans from logical propositions about the world, which would be delivered by an unspecified perceptual system. The typical strategy for finding a plan was to conduct a heuristic search in a large action space. Brooks made two important criticisms of this approach. First, he noted that perception was a major problem in itself, and it was naïve to assume that the "vision guys in white hats down the corridor" would be able to obtain the necessary world descriptions [9]. Second, he argued that search was not the right tool for intelligence. As a counter-point to the symbol system hypothesis, Brooks offered his own *physical grounding hypothesis*: "to build a system that is intelligent it is necessary to have its representations grounded in the physical world" [9]. This view motivated Brooks' research into embodied agents, i.e. real robots operating in the real world.

The ideas of Brooks were pursued vigorously by later researchers [3–5, 10]. An important part of this work is the development of a set of design principles to guide the construction of embodied agents [3]. The principles were obtained through extensive experience with robotic systems and from detailed study of biological organisms. These ideas were primarily targeted toward the physical construction of robots, but they have important implications for the design of learning algorithms as well. In particular, the *Complete Agent Principle* instructs designers to build agents that are "autonomous, self-sufficient, embodied, and situated". Another important idea is the *Principle of Information Self-Structuring*, which states that the agent should take advantage of statistical regularities induced by body-environment interactions, and should actively attempt to seek out such regularities [5, 10, 11]. As we discuss below, it is difficult to reconcile these principles with the current formulation of statistical learning.

3 The Setting of the Embodied Learning Problem

3.1 Two Types of Learning

To motivate the following discussion, we postulate a rough separation of learning into two types: perceptual and behavioral. The former allows the agent to understand the world, while the latter guides the agent's choice of actions. A necessary component of behavioral learning is reinforcement. Agents are assumed to receive a reward signal from the environment that instructs them to behave in an adaptive way: actions that produce positive rewards are strengthened, while actions that produce negative rewards are weakened. Reinforcement learning is an active area of research [12], and the fundamental principles are well understood.

Our view is that reinforcement is sufficient to explain behavioral learning but not perceptual learning. Simply stated, the information from the reinforcement signal is not sufficient to determine the huge complexity of the brain, which has on the order of 10^{12} synapses. For example, it is difficult to believe that differential reinforcement can be used to tune the synaptic weights in the lower levels of the visual cortex.

One can imagine constructing a Complete Agent and equipping it with two learning mechanisms: one for behavioral learning, and one for perceptual learning. For these two learning tasks, one might reasonably choose a reinforcement learning algorithm for the behavioral component, and a CStL algorithm for the perceptual component. The problem with this approach is that the current formulation of statistical learning is not well suited to the type of data encountered by robots. We now outline the setting of the problem of EStL, and contrast it to the CStL formulation. The constrast is summarized in Table 1.

Table 1. Summary of differences between problem formulation in Embodied Statistical Learning (EStL) and the current formulation of statistical learning (CStL).

Aspect of problem	EStL	CStL	Section
Form of data	stream	discrete samples	3.2
Supervisory signal	scarce	frequent	3.3
Volume of data	vast	limited	3.4
Key problem	prediction	recognition	3.6
Agent contribution	actively structures data	passively observes data	3.7

3.2 Data is a Stream

In the real world, for both robots and organisms, data arrives in the form of a stream. No obvious method exists for partitioning the stream into samples that can satisfy the assumptions of CStL. Any such partitioning of the stream will destroy the IID property. If one partitions a stream of visual images into frames, then each frame is strongly dependent on the previous frames.

The data is a stream, but it is not necessary to treat the stream simply as a sequence of bits. For example, if the stream is a sequence of images, it is reasonable to assume that the dimensions of the images are known. Or, if the stream is a sequence of video, audio, and sonar data, then it is reasonable to assume knowledge of which bits correspond to each sensory modality. In return for giving up the assumption of discrete sample data we get a “consolation prize”: the temporal structure of the data stream, which can and should be exploited.

3.3 Labels and Reinforcement Events are Scarce

The Complete Agent Principle instructs us to build agents that are “autonomous, self-sufficient, embodied, and situated” [3]. When applied to learning in robotic systems, this principle requires that the agents should learn in an unsupervised or self-supervised way. The amount of supervisory information provided to the agents, in the form of labeled training data and reinforcement signals, should be strictly limited.

In order to perform a pattern recognition task in CStL, one typically assumes a set of data points (e.g., images) and associated labels. Usually there are as many labels as data points. The labeling process may require a substantial amount of human labor, and is often error-prone. For example, in an image annotation task one must label each image with a set of words describing the objects and activities displayed in the image. However, different people might use different sets of words to describe an image. Also, this model of learning is fundamentally limited to the labels: a vision system trained using labeled data to recognize car models will not be able to determine if the car is parked or in use.

Reinforcement learning for perception faces similar limitations. Imagine we have built a humanoid robot and want to train it to fetch coffee in an office environment. The simplest method would be to give the robot a big reward when it arrives with the coffee. However, this simple scheme will require the robot to explore for years before it happens to retrieve the coffee and get the reward. By adding some complexity to the reward signal, we can potentially improve performance. Maybe we give the robot a small reward for obtaining the coffee, and a larger reward for delivering it. We could then go further, defining rewards for entering the hallway, pouring the coffee into the cup, adding sugar and cream, and so on. We could also define negative rewards for spilling the coffee or bumping into people. However, this violates the Complete Agent Principle, because it requires us to provide greater and greater levels of supervision to the robot, in the form of defining complex reward signals.

In the view of EStL we are developing, the agent is able to learn in an unsupervised and autonomous manner. However, this learning should be thought of as *preparation*, so that when supervisory information arrives, the agent can adapt to it as quickly as possible. When the user gives the coffee fetching robot a command, it should not have to perform a lengthy learning process; it should already know enough about the world to execute the command. Thus, if the robot is a humanoid, it must already have complex knowledge of grasping, walking, the visual stimuli corresponding to coffee cups, and so on.

3.4 The Stream is Vast

Above we argued that data should be thought of as a stream. We now point out that it is an *enormous* stream. Robotic systems can obtain data from cameras, microphones, laser range finders, odometers, gyroscopes, and many other devices.

One of the great insights of statistical learning theory [6, 7] is that when the data is limited, the model employed must be simple. There are various ways to calculate the model complexity, but the idea is the same. A basic rule of thumb is that the complexity of the model cannot exceed the information content of the data being modeled.

Consider the problem of classifying handwritten digits, which can be thought of as a "typical problem" of CStL. We wish to learn a rule that gives a good estimate of the probability distribution $p(Y|X)$ where Y is the label and X is the image. In this case the data being modeled is the set of labels, each of which has an information content of $\log_2 10 \approx 3.2$ bits. Assuming there are 10000 samples, the information content of the entire set of labels is about 32000 bits. Thus we cannot use models that have complexity of greater than 32000 bits. The essence of learning in this low-data regime is to find low complexity models that have high explanatory power. The success of the Support Vector Machines can be attributed to the fact that only a small number of parameters corresponding to the support vectors need to be specified, and the data is separated using the optimal separating hyperplane.

The data available to robots is enormously more vast than the 32000 bits available in the handwritten digit recognition problem. This multiple order of magnitude difference means that the learning problem must be thought about in an entirely new way. In particular, the vast amount of data available justifies the use of highly complex models. To sum up, the basic problem of CStL is: how can one generalize well from a limited amount of data? In contrast, the basic problem of EStL is: how can one efficiently exploit the huge amount of data to build a complex model of the complex world?

3.5 Fast is better than Slow

This point is a combination of three distinct ideas, all of which emphasize speed in different ways. The first idea is that an agent must react rapidly to supervisory information when it appears, as was illustrated by the example of the coffee fetching robot. Similarly, when considering biological situations, strong negative reinforcement signals often relate to life-threatening events (e.g. a rabbit eating a poisonous plant), so the agent must adapt to those signals rapidly. To allow the behavioral learning component to adapt quickly, the perceptual learning component must provide it with meaningful abstractions.

The second reason for emphasizing speed is a consequence of our emphasis on learning vast data. If the computational architecture cannot process the data efficiently, it will choke on the vast size of the stream. In CStL, the learning bottleneck is the limited amount of data; in EStL it will likely be the computational complexity of learning.

The third idea is that the learning process should be primarily online. Learning should begin immediately once the sensor data stream starts flowing, and should proceed in a continuous fashion thereafter. Ideally, at each step the learning machine should update itself to reflect the new piece of data that has arrived.

3.6 Prediction is Critical

The claim of this section is that the ability to predict is necessary and sufficient for intelligent behavior in the sense of optimizing future reward. To see that prediction is sufficient for reward-optimizing behavior, consider the following reinforcement learning strategy. We assume that the agent has experienced a large amount of data with sensor, motor, and reward components. It has built a model which interleaves these data types, allowing it to predict the future reward from the sensor stimulation and motor actions. Then it predicts the future reward given the current stimulus and a variety of action plans, and chooses the plan corresponding to the highest predicted reward.

To see that prediction is necessary, notice that real agents must be able to predict that an action like jumping off a cliff will be harmful without actually experimenting with it. This is fairly obvious, but standard reinforcement learning algorithms do not provide a mechanism to avoid bad states without actually visiting those states.

Note that other problems of classification, recognition, and so forth can be thought of as subproblems of prediction. If one can recognize the numbers and letters written on a business card, one can predict the identity of the person who will answer the phone when the number is dialled.

Because of its relationship to reward-optimization and classification, prediction can be thought of as a fundamental cognitive task. Thus, if a powerful and general purpose prediction method can be achieved, it will bring us much closer to the goal of intelligent machines. The critical role of prediction was recently articulated by Hawkins [13].

3.7 Agents Influence their Own Learning

An agent learns by adapting to the vast stream of data entering its experience. However, some types of data are better suited to the learning process than others. For example, it is probably not very useful to observe visual data from a television tuned to a dead channel. On the other hand, data that has a certain kind of statistical regularity may be especially helpful in guiding the learning process. We say that this type of data is *structured*.

The Principle of Information Self-Structuring discussed in Section 2.2 states that an agent should actively attempt to induce structure in the data entering its experience. If this can be done successfully, the learning machine will be able to adapt more rapidly to the environment.

One mechanism of information self-structuring is the idea of intrinsic rewards for learning, referred to as the "autotelic drive" by Steels [11]. In order to implement an autotelic drive, the learning machine reports a signal that characterizes the degree of information structure in the incoming stimuli. It should also reflect the extent to which the information structure is useful in *improving* the performance of the learning machine: even if a certain pattern is highly structured, it may not be useful to observe it repeatedly after it has been thoroughly learned. This intrinsic reward is combined with external reward to guide behavior.

Another mechanism of information self-structuring is morphological computation, which is the idea that the body can act as a computational device to reduce the cognitive burden on the brain [14]. For example, it can be shown that a fly's eye is morphologically suited to the problem of detecting motion, because of the curvature of the lens [15]. The lens preprocesses the incoming sensory data in such a way as to simplify the computational problem of motion detection.

The Principle of Information Self-Structuring does not fit easily into the CStL paradigm. With the important exception of research in active learning [16], most CStL algorithms assume that the agent itself plays no part in the selection of data points used for training. This is required because if the robot uses information from the first $\frac{N}{2}$ data points to decide how to select the next $\frac{N}{2}$ samples, the IID assumption breaks down. Thus, the CStL theory does not provide insight regarding how to implement an autotelic drive, or how to design agent bodies to facilitate fast learning.

4 Synthesis: Adaptive Compression

Historically, there have been two main paradigms in statistical learning. The first is that of learning as induction, described by the Vapnik quote above. The second is the view of learning as compression, which has its roots in the idea of Minimum Description Length modeling [7]. There is a deep relationship between induction and compression [6, 7].

In the above discussion, we described what we consider to be the proper inputs to the learning algorithm. We are now faced with the question of what the learning algorithm should actually *do* with the input. We claim that the only meaningful learning process that can be performed on the basis of a vast stream of unlabeled data is adaptive compression. Specifically, the learning algorithm should incrementally update a statistical model so as to reduce the bit rate per unit time required to represent the incoming data stream. This view connects directly to the idea of *redundancy reduction* which has been proposed as a fundamental principle explaining the function of the cortex [17].

Compared to the compression view, the induction view may seem more attractive for practical reasons. A program that can determine if a face is present in an image may seem more useful than a program that can compress images with faces in them. Thus, to further justify the goal of compression, we propose the following hypothesis: *statistical models obtained through the adaptive compression process will be useful for other applications.* There is a variety of evidence for this hypothesis. In recent work by Hinton *et al.*, it is demonstrated that building a generative model of handwritten digit images is useful in recognizing their labels [2]. Also, in the field of statistical natural language processing, improvements in the language model immediately yield improvements in applications such as speech recognition and machine translation [18]. In both cases, the model is obtained by finding a set of parameters that minimizes the log-likelihood of the original data (text or digit images); this process is basically equivalent to compression.

Note also the strong link between compression and prediction. If one can predict a data stream, then one can compress it. Thus while we use the compression rate for comparison purposes because it is a hard number, what we are really measuring is an algorithm's ability to *predict*.

It is important to note that this view is agnostic with regard to the choice of computational approach (e.g. dynamical systems or physical symbol systems) underlying the learning process. Given a computational model, it is easy to construct a compression algorithm on top of it. Thus the compression rate can be used by advocates of various perspectives on cognition to provide strong quantitative evidence for their views. One simply constructs a compression algorithm inspired by a particular idea about cognition and applies it to some large dataset (ideally, a benchmark dataset). If the new algorithm achieves a significant reduction in compressed data size, this provides strong quantitative evidence for the cognitive model. Thus, the view of learning as compression supports a rigorous methodology, and we consider this to be one of the major arguments in favor of it. This rationale for using compression rates to quantify progress in artificial intelligence research was recently articulated by Mahoney [19].

As an example of how the compression methodology can benefit embodied agents research, consider the work of Tani and Nolfi [20], which describes a method for hierarchical learning of different categories of sensory-motor data generated by a mobile robot. The authors show that the recurrent neural network modules self-organize such that each module becomes an expert at one type of data. This is an interesting result, but it is difficult to compare the method to other possible techniques. If the paper reported the compression rate achieved by the system on the sensory-motor data, it would be a much more powerful vindication of the method. In work using a similar experimental setup, but a very different modeling scheme, we showed how the compression rate can be used as a performance measure in cases where there is no obvious task to perform (i.e., the robot is simply exploring without a specific goal) [21]. The important result here is that the statistical model obtained in this process is useful for other tasks such as localization, thus supporting the hypothesis given above. However, our modeling method is fairly simplistic; it is likely that other methods (such as the one proposed by Tani and Nolfi) will provide better performance. The compression rate should allow us to select the best general method.

5 Conclusion

The goal of Embodied Statistical Learning is to fuse together the strong mathematical theory of statistical learning with the design principles of Embodied AI. This requires a new setting for the learning problem, because of the mismatch between the type of data available to embodied agents, and the type of data assumed by the current theories. In the new formulation, the input data is a vast unlabeled stream which is actively structured by the agent. We argued that the only meaningful learning process that can be applied to a vast unlabeled data stream is adaptive compression. Compression is equivalent to prediction, and allows for rigorous comparisons of results. We also hypothesize that the statistical model obtained through compression will be useful for other applications.

Compared to CStL, we consider EStL to be a more realistic setting of the learning problem. It may also be an *easier* setting, for the following reasons. First, the agent can exploit the temporal structure of the data stream. Second, the agent can perform information self-structuring. But the most important reason is that the amount of data available is enormous. The exploitation of this vast data resource may allow us to construct models of complexity comparable to the human brain.

References

1. Vapnik, V., Golowich, S.E., Smola, A.J.: Support vector method for function approximation, regression estimation and signal processing. In: Advances in Neural Information Processing Systems. (1996) 281–287
2. Hinton, G., Osindero, S., Teh, Y.: A fast learning algorithm for deep belief nets. Neural Computation **18** (2006) 1527–1554
3. Pfeifer, R., Scheier, C.: Understanding Intelligence. Cambridge, MA: MIT Press (1999)
4. Pfeifer, R., Lungarella, M., Iida, F.: Self-organization, embodiment, and biologically inspired robotics. Science **318** (2007) 1088–1093
5. Kuniyoshi, Y., Yorozu, Y., Suzuki, S., Sangawa, S., Ohmura, Y., Terada, K., Nagakubo, A.: Emergence and development of embodied cognition: A constructivist approach using robots. In von Hofsten, C., Rosander, K., eds.: From Action to Cognition. Volume 164 of Progress in Brain Research. Elsevier (2007) 435–455
6. Vapnik, V.: The Nature of Statistical Learning Theory. Springer (1998)
7. Rissanen, J.: Modeling by shortest data description. Automatica **14** (1978) 465–471
8. Brooks, R.A.: Intelligence without representation. Artificial Intelligence **47** (1991) 139–160
9. Brooks, R.A.: Elephants don't play chess. Robotics and Autonomous Systems **6** (1990) 3–15
10. Lungarella, M., Sporns, O.: Information self-structuring: Key principle for learning and development. In: International Conference on Development and Learning. (2005)
11. Steels, L.: The autotelic principle. In: Embodied Artificial Intelligence. (2003) 231–242
12. Sutton, R., Barto, A.: Reinforcement Learning: An Introduction. Cambridge, MA: MIT Press (1997)
13. Hawkins, J., Blakeslee, S.: On Intelligence. New York: Times Books (2004)
14. Pfeifer, R., Iida, F., Gomez, G.: Morphological computation for adaptive behavior and cognition. International Congress Series **1291** (2006) 22–29
15. Franceschini, N., Pichon, J., Blanes, C.: From insect vision to robot vision. Philosophical Transactions: Biological Sciences **337**(1281) (1992) 283–294
16. Tong, S., Koller, D.: Support vector machine active learning with applications to text classification. Journal of Machine Learning Research **2** (2001) 45–66
17. Atick, J.: Could information theory provide an ecological theory of sensory processing? Network **3** (1992) 213–251
18. Rosenfeld, R.: A maximum entropy approach to adaptive statistical language modeling. Computer, Speech and Language **10** (1996) 187– 228
19. Mahoney, M.: Rationale for a large text compression benchmark. `http://www.cs.fit.edu/~mmahoney/compression/rationale.html` (2006)
20. Tani, J., Nolfi, S.: Learning to perceive the world as articulated: an approach for hierarchical learning in sensory-motor systems. Neural Networks **12**(7-8) (1999) 1131–1141
21. Burfoot, D., Kuniyoshi, Y.: Maximum entropy statistical modeling of sensor data from robotic systems. In: Proceedings of the Symposium on Language and Robots. (2007) 107–116