Towards Learning 2.0

Etienne Barnard
Council for Scientific and Industrial Research
South Africa
Email: ebarnard@csir.co.za

Brigitte Palensky
Austria
Email: brit@mond.at

Peter Palensky
University of Pretoria
South Africa
Email: palensky@ieee.org

Dietmar Bruckner
Vienna University of Technology
Austria
Email: bruckner@ict.tuwien.ac.at

***Abstract*—Learning certainly qualifies as one of the core issues of artificial intelligence (AI). During the years, it has gained – and subsequently lost – popularity in the research community. After a historical perspective on the rise and fall of learning research in AI, some of the limitations of current learning systems are reviewed, followed by a presentation of various responses about how to overcome them. A special focus is given on one of the responses, the attempt to draw lessons from a detailed study of evolutionary and developmental processes and stages of learning in nature, in particular in human beings. From this, a number of principles for machine learning are inferred. A key aspect seems to be that learning should be cumulative to compensate for the exponential growth in learning complexity.**

I. Introduction

> *If you invent a breakthrough in artificial intelligence, so machines can learn, that is worth 10 Microsofts.* Bill Gates, quoted in New York Times, March 3, 2004.

> *Only a small community has concentrated on general intelligence. No one has tried to make a thinking machine and then teach it chess [...] We have got to get back to the deepest questions of AI and general intelligence and quit wasting time on little projects that don't contribute to the main goal.* Marvin Minsky, quoted in [33].

Information Technology dips into the Artificial intelligence (AI) toolbox whenever one of the following situations occurs:

- A problem is too complex to model it analytically
- A solution is too complex to implement it synthetically
- You need a fancy sales term in your brochure

AI has certainly created a set of very *useful* tools with remarkable capabilities. Successful enterprises like Amazon and Google rely on filters, feature extractors and other AI mechanisms for their business intelligence. However, these achievements must be calibrated against the early expectations of "machines ... capable ... of doing any work a man can do" [32, p. 96] or "a machine with the general intelligence of an average human being" (M. Minsky, in a contested quote in Life magazine in 1970).

What has happened to the dream of truly intelligent artifacts? What has happened to the goal of machines that would have virtually no cognitive limits, with human-level performance a mere milepost along the road of increasingly intelligent automata? From one perspective, those dreams have turned out to be irrelevant to many of the practical applications of AI that have impacted society so widely. Nevertheless, to many the achievement of "Artificial General Intelligence" (AGI) continues to be the touchstone for assessing progress in the field. For example, the so-called "AI Singularity" is expected to happen after an AI reaches beyond the capabilities of our own "wetware" [35]. At that point, an artificial system will be able to create even more intelligent machines, leading to even further automatic enhancements, and thus to a cycle of ever more intelligent systems. Within short time, in this vision, we will be surrounded by highly sophisticated intelligent automata, far beyond the comprehension of mere mortals. (The AI Singularity has evangelists, critics and agnostics; the most prominent opinions can be reviewed in the June 2008 issue of the IEEE Spectrum.)

Most researchers who subscribe to this theory base their expectations on the exponential growth of technology, which in turn is one of the points of criticism leveled at the singularity hypothesis. Some imagine that the anticipated worldwide network of augmented objects – small computers, embedded into everyday objects – will spontaneously form a collective consciousness (which might eventually include humans), whereas others expect supercomputers to create the first AI singularity.

Thus, the continued interest in AGI has not produced any consensus on the most promising research directions to be pursued. In the current contribution, we take our cue from the fact that human development is driven by learning. As pointed out by Brooks [5], two year old children can reliably discriminate objects, four year olds can carry out sensible conversations, six year olds have remarkable manual dexterity in operating objects and eight year olds already have sophisticated social abilities and a theory of mind of others. On the other hand, extremely sophisticated and ambitious attempts [18] to encode common-sense knowledge in a pre-arranged, unlearned fashion have not been able to approach similar levels of performance. We believe that the balance of the evidence argues for the importance of learning in intelligent systems: without learning capabilities, machines will never achieve general intelligence.

Below, we first provide a historical perspective on the cycles of excitement and disappointment that have characterized the history of learning research. Thereafter, we review some of the limitations of current learning systems, in order to understand the reasons why such systems have not lived up to the lofty expectations of successive generations of researchers. We then investigate various responses to this situation, and summarize a number of principles that can be inferred from a study of human learning.

II. The time of learning has come, and passed, and come...

AI has come a long way since the Dartmouth conference in 1956 – today, AI is used on a daily basis by millions of people, and large sections of modern economies rely on capabilities that have sprung from AI. However, this progress has not been monotonic: there have been waves of excitement, interspersed with periods of widespread disappointment in what has been achieved.

Learning has had an important role in several acts of this play. In the original proposal for the Dartmouth conference, learning was featured as a core aspect of AI [20], and it was not long afterwards that algorithms such as the perceptron algorithm of Rosenblatt [30] seemed to create a basis for general-purpose learning. Rosenblatt himself was optimistic that this work would lead towards "the fundamental laws of organization which are common to all information handling systems, machines and men included". However, during the 1960s most progress in AI arose from work that did not contain a learning element. Thus, topics such as search, the design of heuristics and knowledge representation became the core activities in AI, and the analysis of perceptron limitations by Minsky and Papert [21] was interpreted by many as a rationale for this state of affairs.

A new wave of activity in automated learning was introduced by the "connectionist revival", stimulated primarily by the work of Hopfield [16] and Rumelhart and collaborators [31]. This work not only overcame the fundamental limitations pointed out by Minsky and Papert (based on the observation that only linear discriminations were possible with the original perceptron learning algorithm); it also found practical application in diverse areas such as language processing, credit scoring and image recognition. It again seemed possible that a general-purpose learning system could be developed, and in graduate schools around the world thoughts similar to those of Rosenblatt were aired. Given sufficient computational resources and raw data, it was hoped, any desired behavior would eventually be learned. As happened to perceptrons, this hope eventually disappeared as both theoretical and experimental evidence mounted against it. On the experimental side, it became clear that reasonable success could only be obtained with carefully pre-processed data [2], which in effect limited the complexity of the learning task. Theoretically, several results were obtained (for example, [4]) indicating that the general learning problem for multilayer perceptrons is computationally intractable.

Although alternative learning approaches have been developed in the wake of the connectionist era (for example, maximum margin classifiers and graphical models), it is safe to say that the main focus in AI has again shifted away from learning methods. No single dominant paradigm is currently in operation, but embodied systems, intelligent agents and a renewed focus on biomimetic systems have – in certain cases – shifted the attention to aspects of artificial intelligence that are not directly aligned with learning (although, as we discuss in later in this paper, biomimesis and learning may also align very well).

III. The limitations of learning, ca 2008

Various forms of learning algorithms certainly play an important role in some of the successes that have been achieved with AI. For example, in speech processing, statistical learning algorithms are indispensable to the development of systems that are used by millions of people every day; credit scoring with neural networks is only possible because of their trainability, and collaborative filtering is by its very nature a learning enterprise. These successes of optimization-based learning systems seem like a convincing rebuttal of the pessimism displayed in statements such as the following:

"In the early years of cybernetics, everybody understood that hill-climbing was always available for working easy problems, but that it almost always became impractical for problems of larger sizes and complexities ..." [22, pp. 260-261]

However, a deeper analysis reveals that the intuition expressed by Minsky and Papert remains valid. As mentioned above, successful learning in each of these cases requires that the inputs to the learning system, and the outputs expected of it, be chosen very carefully to limit the complexity of the mapping that it needs to compute. If this is not done, learning is generally not successful: typically, the number of inferior local extrema encountered by an optimizing algorithm becomes unmanageably large.

This dichotomy is – to some extent – explained by theoretical results such as that of Gold [14] and of Blum et al. [4], which demonstrated that several realistic learning problems were NP complete, and could therefore not be solved for large problem spaces. Although such proofs differ greatly in their technical details, the fundamental cause for intractability can generally be traced back to the exponential growth in the number of underlying structures that can conceivably give rise to an increasingly large set of observations.

A distinct set of objections against general-purpose learning have been constructed around the so-called "no free lunch theorem" for learning [12]. That theorem states that, under certain apparently general conditions, no learning algorithm can reliably outperform any other [36]. This is interpreted to mean that prior, unlearned, information plays a crucial role in cases where learning does indeed succeed, thus casting doubts on the possibility of a completely general learning algorithm. However, we have recently argued [1] that the apparently innocuous conditions of the NFL theorem are in fact highly unnatural; consequently, we doubt the relevance of this result.

It is nevertheless true that current learning algorithms are severely deficient in comparison to the highly flexible learning capabilities that are observed in nature, and the conflict between this observation and the importance of learning for which we have argued in the introduction poses an important challenge. Possible responses to this challenge are investigated next.

IV. Maps of the forking path

One may argue that the limited results achieved by learning algorithms (after intensive research) demonstrates that our naive expectations of the power of general learning algorithms were exaggerated (much as special relativity demonstrated that naive expectations about the maximum velocity that a physical body can obtain were not achievable). Such lowered expectations are indeed contained in a number of thoughtful responses to this "learning paradox". Brooks, for example, argued [6] that it neither feasible nor necessary to construct deep representations, or models, of the external world. (As we will argue below, the construction of such representations lies at the core of the learning challenge.) I, Brooks argues for systems that use the world as its own model – that is, systems that repeatedly update their own behavior based on the state of the world (as obtained from direct observation). If necessary, the system can purposefully modify aspects of the external world to keep track of its internal changes.

Other proposals from the perspective of limited learning power include the modular model of the world, as described (for example) in [26]. Here, the idea is that evolution has provided brains with several task-specific modules that can function well without much learning; Development of skills and abilities, in this view, is more a matter of parameter setting than learning from scratch. Even more radically, Baum [3] argues for the encoding of much of the basic information that underlies adaptive performance directly in DNA.

Although each of these theories has gathered significant support, they continue to be minority views. The most convincing counter-argument is that our ability to build space ships, or digital telephone networks, requires both substantial internal representations and an ability to go very far beyond whatever abilities or knowledge may be innate. A sophisticated learning theory would therefore be necessary even if such modules do in fact exist.

An alternative argument (which in some form has been around since the earliest days of AI) attributes our current limitations to insufficient computational resources. Roughly speaking, the argument here is that our algorithms are on the right track, but we need to harness vastly more computational power to achieve human-like intelligence. Thus, for example, the OpenCog project [15] aims to develop a software platform that makes it feasible to run extremely large simulations of heterogeneous, interactive "cognitive routines".

A related, though distinct, appeal to the transforming power of massive computational resources comes from the drive to ever greater biological realism. It is clear, for example, that current artificial neural networks are based on grossly simplified abstractions of the behaviors of biological neurons. It is quite plausible that the essential capabilities of real neurons are lost in this abstraction, and authors such as Kurzweil [17] and Markram [19] have therefore argued for the need to deploy vast computational power in order to create detailed, molecule-by-molecule, simulations of biological neural systems. The BlueBrain project is an ambitious step in this direction. However, it is still hotly contested whether such brute-force simulation can be expected to produce intelligent behavior in the absence of theoretical insights into the functioning of the systems being modeled. In particular, the question remains whether even these highly sophisticated simulations really capture the key ingredients that make intelligent behavior (and learning) possible. Some authors [10] claim that even the most detailed simulation or computation of the physical structure of the brain cannot provide insights into the functional structure of thinking including learning, until many levels of abstraction dealing also with feedbacks inbetween are not included into the model.

Finally, the hope still persists that algorithmic research will uncover principles that will make general-purpose learning possible – despite everything that has happened to date. Progress in fields such as graphical modeling continues to promise that a resolution of the theoretical difficulties of learning theory can be found, possibly through insights on the automatic decomposition of learning spaces, or from a more detailed study of those extremely capable learning systems that occur in nature, as we discuss next.

V. Learning from Nature

Instead of trying to achieve greater biological realism regarding the material substrate of cognitive capacities (as e.g. in the BlueBrain project), there is also the possibility of gaining insight about how to construct powerful internal representations by having a closer look at how this is achieved by human beings during their years of infancy and childhood. A model of the changes of cognitive representations during these early years that derives some of its fundamental aspects from a comparative analysis of developmental and evolutionary processes and stages of human learning, cognition, and language is presented by K. Nelson in [24].

From a general evolutionary perspective, according to Plotkin [27], the following four levels of knowledge acquisition can be distinguished:

1) Genetic programs constituting organismic adaptations to environmental conditions that vary only very slowly, that is, over generations.
2) Variable developmental programs constituting adaptations to environmental conditions that vary during the lifespan of an organism. Variations can occur in morphological development, but also in behavioral development. Possible processes that produce variation include acceleration, delay, or heterochrony across different components of a system.
3) Individual learning processes constituting the most general adaptive mode of organisms. Possible processes include habituation, classical or operant conditioning.

4) Group level adaptations built on information sharing communication systems. Various social learning mechanisms [37] are included, for example true imitation.

The whole picture the framework delivers is that of a hierarchy of nested processes, built upon each other and inherently intertwined. Plotkin writes [27, p. 85]: "Tabula rasa learning is impossible because the third level evolved out of a failure of the first and second levels to deal with changes occurring above a certain frequency .. it is rooted within and constrained by those more fundamental levels of the hierarchy".

According to D. Campell [9], a fifth level of knowledge accumulation can be added:

5) Adaptations achievable via cultural representations and semiotic cumulations, examples including oral traditions, or libraries.

In [34], M. Tomasello argues that the level of cultural learning leads to a dramatic accumulation of knowledge where individuals gain access to the collective previous achievements of their social group. Furthermore, the emergence/invention of human language, in its essence a complex symbolic form of communication, is maintained to be strongly related to the cultural level.

In [11] a more detailed proposal of how culture and cognition are mutually constitutive and how language fits into the picture is presented. In this model, focusing on the evolution of the human mind from the basic primate mind, Donald suggests a series of three major stages of adaptations, each of which comes with the emergence of a new representational system:

1) Episodic culture. This stage characterizes the general primate mind and is based on event perception, "the ability to perceive complex, usually moving, clusters and patterns of stimuli as unit" [11, p. 153]. Cognition on this stage is situation-bound, concrete, and unreflective.
2) Mimetic culture. "Mimesis" is defined as imitation to a higher purpose. Requirements are a general capacity to mime, intentional representations, a public communicative system, differentiation of reference, autocued rehearsal, and the ability to model an unlimited number of objects. Mimetic skills bring about the possibility of sharing knowledge without the necessity of every group member reinventing it, and of modeling social customs and structure ("rituals").
3) Mythic culture. On the basis of primitive semiosis existent in mimetic cognition, language/speech is built as a discourse mechanism, serving social purposes and the integration of extended thematic parts.

The invention of human language in the mythic stage, bringing existing pre-symbolic mental models under symbolic control, is claimed to have led to a radical change in the functioning of the human mind. With language, components of event memories can be independently accessed and entered into new, not necessarily situation-bound, constructions. Equally important, language opens a window into other minds allowing, among other things, experienced individuals to make implicit relations manifest in learning situations with inexperienced individuals. As a collectively established, representational and mediational system, language amplifies thinking and learning beyond personal, individual experience.

What implications may be drawn for machine learning from the presented biological account of cognitive evolution and development? Before a list of suggestions is presented, a general remark shall be made. In nature, the vast majority of learning is carefully specialised to task [13]. This may be due to the complexity issues of the learning problem as revealed by mathematical analysis. Even if individuals of a species possess some general learning capacity (like human beings), it is in their interest to learn knowledge from similar agents (via social and cultural learning) because the probability of stumbling across a useful piece of knowledge by themselves during their lifetime might be very low. So, as in nature, the general learning problem in AI may not be solved by one kind of general-purpose learning algorithm, but by a mixture of algorithms, each of them potentially working on different levels (analog to classical and operant conditioning, chaining, concept learning, principle learning, imitation, logical deduction, etc.), and each of them providing complexity-reducing mechanisms appropriate to its level of working.

What follows is a list of speculations about the lessons to be learned from nature regarding learning and knowledge acquisition in AI:

1) Representations should be hierarchical, starting with a basic structure that can efficiently process huge amounts of various kinds of data points, and ending with higher-level structures that apply to many different domains of knowledge, that can express temporal, causal, conditional, and intentional relations, that support generalizations of patterns on a highly abstract level, and that allow explicit communication with other systems (machines and human beings) so that more and more different programs (and with them, also their different domains of knowledge) can be integrated into one big knowledgeable program.
2) Instead of focusing on objects, event representations should be used as basic building blocks of cognition. The focus should be on guiding (own) action and predicting activities (of others). What is experienced should become a function of the current concerns/tasks/purposes.
3) For different purposes, different representations can be used [7].
4) The higher-level representational structures referred to in 1) may require symbolic representations as powerful as human language [29].
5) Learning need not start from a completely blank state. Programs should be tailored as efficiently as possible to their respective task/field of application. Heuristic knowledge in-built by the programmer could be compared to genetic knowledge, both are a-priori available.
6) Regardless of 5), a phase in which the perceptual and conceptual sub-systems of a program are "tuned"

to the specifics of its environment/field of application seems indispensable (analog to infancy). This requires to provide mechanisms for representational change so that reprogramming based on experience can take place. The environment includes physical, social, and cultural aspects.

7) Experience need not/should not only be gathered in a passive way. Like infants, programs in their tuning phase could be explicitly instructed. They could also be enabled to explicitly ask questions to clarify ambivalent or unclear aspects of a learning problem.

Some of the aspects above are discussed in more detail in [25]. The cognitive architecture presented there uses event representations as core data structures.

VI. Outlook and Conclusion

Even though much progress has been made in the design of learning machines during the past five decades, there is little doubt that learning has not fulfilled its early promise. There is nevertheless strong evidence that a sophisticated learning capability is crucial to general intelligence. Several competing perspectives on this conundrum exist, and it is therefore interesting to speculate whether there are near-term results which may suggest the eventual outcome of these debates (whereas, for example, the Turing test can be viewed as a distant goal). We believe that a reasonable hypothesis for such a near-term test focuses on the extent to which learning is cumulative. In humans (and related species), such cumulative learning is highly salient – it is commonplace that an expert in any domain is able to learn a particular fact or procedure in that domain much more quickly than a novice. (Childhood learning is an extreme example of this phenomenon.) However, in current learning algorithms, such a cumulative learning ability is virtually non-existent. Perhaps the only example of improved learning with time is found when adaptive learning rates are used in iterative learning schemes; however, such benefit is severely limited and tends to disappear after a relatively short period.

It is intuitively plausible that cumulative learning can compensate for the exponential growth in learning complexity described in Section III: if each cumulative step is engenders a multiplicative improvement, an ideal cumulative learning strategy will have a compensating exponential growth in capabilities. Of course, any process characterized by exponential growth will eventually saturate, so the expectation is that the accumulative phase of learning will eventually tail off – a phenomenon that is certainly evident in human learning.

Moores law has been driving progress in information technology for more than three decades. If our speculations are correct, another form of exponential improvement is likely to take on that role when fundamental progress in learning is made. That will certainly qualify as "Learning 2.0"!

References

[1] E. Barnard, "Learnability and the "No Free Lunch" paradox", Submitted for publication, 2008

[2] E. Barnard, D. Casasent, "Invariance and neural nets", IEEE Transactions on Neural Networks, Vol. 2, Nr. 5, pp. 498-508, 1991.

[3] E.B. Baum, "What is thought", MIT Press, 2004.

[4] A. Blum , R.L. Rivest, "Training a 3-node neural network is NP-complete", Proceedings of the first annual workshop on Computational learning theory, pp. 9-18, August 03-05, 1988, MIT, Cambridge, Massachusetts, United States, 1988.

[5] R.A. Brooks, "I, Rodney Brooks, Am a Robot", IEEE Spectrum 06.08, 2008.

[6] R.A. Brooks, "Intelligence without representation", Artificial Intelligence 47, 139-159, 1991.

[7] D. Bruckner, "Probabilistic Models in Building Automation, Recognizing Scenarios with Statistical Methods", Dissertation Thesis, Institute of Computer Technology, 2007.

[8] D. Campbell, "Evolutionary epistemology", in H.C. Plotkin (Ed.), Learning, development and culture, pp. 73-107, New York: Wiley, 1982/1974.

[9] D. Campbell, "Levels of organization, downward causation, and the selection-theory approach to evolutionary epistemology", in G. Greenberg & E. Tobach (Eds.), Theories of the evolution of knowing, pp. 1-18, Hillsdale NJ: Erlbaum, 1990.

[10] Dietrich, D., Fodor, G, Zucker, G., Bruckner, D. (eds.): Simulating the Mind. Springer (2008)

[11] M. Donald, "Origins of the modern mind: Three stages in the evolution of culture and cognition", Cambridge, MA: Harvard University Press, 1991.

[12] R.O. Duda, P.E. Hart, D.G. Stork, "Pattern Classification", New York: Wiley-Interscience, 2001.

[13] C. Gallistel, A.L. Brown, S. Carey, R. Gelman, and F.C. Keil, "Lessons from animal learning for the study of cognitive development", in S. Carey and R. Gelman (editors), The Epigenesis of Mind, pp. 3-36, Lawrence Erlbaum, Hillsdale, NJ, 1991.

[14] E.M. Gold, "Language Identification in the Limit." Information and Control 10:447-474. 1967.

[15] D. Hart, B. Goertzel, "OpenCog: A Software Framework for Integrative Artificial General Intelligence", available online at www.agiri.org/OpenCog_AGI-08.pdf

[16] J.J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities", Proceedings of the National Academy of Sciences of the USA, vol. 79 no. 8 pp. 2554-2558, 1982.

[17] R. Kurzweil, "The Singularity Is Near: When Humans Transcend Biology" Viking, 2005.

[18] D.B. Lenat , R.V. Guha, K. Pittman, D. Pratt and M. Shepherd, "Cyc: toward programs with common sense", Communications of the ACM, Volume 33, Number 8, pp. 30-49, 1990.

[19] H. Markram, "The Blue Brain Project", Nature Reviews Neuroscience 7, 153-160, 2006.

[20] J. McCarthy, M. Minsky, N. Rochester & C. Shannon. "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence", unpublished document, 1955. Available at http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html

[21] M.L. Minsky and S.A. Papert, "Perceptrons: An Introduction to Computational Geometry", MIT Press, Cambridge MA, 1969.

[22] M.L. Minsky and S.A. Papert, "Perceptrons – expanded edition: An Introduction to Computational Geometry", Cambridge, MA: MIT Press, 1988.

[23] K. Nelson, "Event knowledge: Structure and function in development", Hillsdale, NJ: Erlbaum, 1986.

[24] K. Nelson, "Language in Cognitive Development", Cambridge University Press, 1996.

[25] B. Palensky, "From Neuro-Psychoanalysis to Cognitive and Affective Automation Systems", VDM Verlag, Saarbruecken, 2008

[26] S. Pinker, "The Language Instinct", New York: Morrow, 1994.

[27] H.C. Plotkin, "An evolutionary epistemological approach to the evolution of intelligence". In H.J. Jerison & I. Jerison (Eds.), Intelligence and evolutionary biology, pp. 73-92, New York: Springer Verlag, 1988.

[28] H.C. Plotkin, "Darwin machines and the nature of knowledge", Cambridge, MA: Harvard University Press, 1993.

[29] G. Pratl, P. Palensky,"Project ARS - The next step towards an intelligent environment", Proceedings of the IEE International Workshop on Intelligent Environments", 2005.

[30] F. Rosenblatt, "The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain", Psychological Review, v65, No. 6, pp. 386-408, 1958.

[31] D.E. Rumelhart, J.L. McClelland, "Parallel distributed processing: explorations in the microstructure of cognition", MIT Press, 1986.

[32] H.A. Simon, "The Shape of Automation for Men and Management", New York: Harper & Row, 1965.

[33] D.G. Stork (editor), "HAL's Legacy: 2001's Computer as Dream and Reality", MIT Press, 1997.

[34] M. Tomasello, "A.C. Kruger, H.H. Ratner, Cultural learning", Behavioral and Brain Sciences, 16, pp. 495-552, 1993.

[35] V. Vinge, "The Coming Technological Singularity", Whole Earth Review, Winter 1993.

[36] D.H. Wolpert, "The Lack of A Priori Distinctions between Learning Algorithms", Neural Computation, 8, pp. 1341-1390, 1996

[37] T.R. Zentall, "Imitation in animals: Evidence, function, and mechanisms", Cybernetics & Systems, 32(1), pp. 53-96, 2001.