Advances in Artificial Intelligence and Machine Learning; Research 1 (2) 136-154 Received 20-07-21; Accepted 05-08-21; Published 20-08-21

Why Learning and Machine Learning Are Different

Janos Vegh VEGH.JANOS@GMAIL.COM
Kalimános BT, Komlóssy u 26, Debrecen, 4032, Hungary.
AdamJozsef Berki BERKI.ADAM@YAHOO.COM
University of Medicine, Pharmacy, Science and Technology of Târgu Mureş, 540142 Târgu Mureş, Romania.

Corresponding Author: Janos Vegh.

Abstract

Machine learning intends to be a biology-mimicking learning method, implemented by means of technical computing. Their technology and methods, however, differ very much; mainly because technological computing is based on the time-unaware classic computing paradigm. Based on the time-aware computing paradigm, the paper discovers the mechanism of biological information storing and learning; furthermore, it explains, why biological and technological information handling and learning are entirely different. The consequences of the huge difference in transmission speed in those computing systems may remain hidden in "toy"-level technological systems but comes to the light in systems having large size and/or mimicking neuronal operations. The biology-mimicking technological operations are in resemblance to the biological operations only when using time-unaware computing paradigm. The difference leads also to the need of introducing "training" mode (with desperately low efficiency) in technological learning, while biological systems have the ability of life-long learning. It is at least misleading to use technological learning methods to complement biological learning studies. The examples show evidence for the effect of transmission time in published experiments.

Keywords: Learning, machine learning,Computing efficiency,Temporal logic, Time-Aware computing, Neuronal computing, Computing paradigm.

1. INTRODUCTION

The ability to learn is a key factor in both the evolution of life and our individual survival. "The broad definition of learning: use present information to adjust a circuit, to improve future performance" needs explanation also [1], what is the "present information", and information at all; furthermore it implies that learning is a temporal process. There are many different definitions what actually 'information' means [2]. In neural science it was observed and theoretically discussed [3] that when using neuronal spikes, "the more precisely spike timing is measured, the more information is gained". However, the conclusion, that timing delivers information, is missing from the information

Citation: Janos Vegh and AdamJozsef Berki. Why Learning and Machine Learning Are Different. Advances in Artificial Intelligence and Machine Learning. 2021;1(2):9.

theory. Information handling is an overly complex process, and it includes coding, decoding, transferring, storing, and retrieving processes, among others.

During learning, the organ uses the past information stored in its internal state variables. However, “we should not seek a special organ for ‘information storage’ — it is stored, as it should be, in every circuit” [1]. The definition also suggests to consider learning (and, because of this: information processing) as a temporal process: the organs have a sensing time, a processing time, a storage access time (despite that “information stored directly at a synapse can be retrieved directly” [1]) and a transfer time (conduction time) to the next computing organ. The general model of computing [4], based on the time-aware computing paradigm [5], correctly describes the general learning process (is underpinned by anatomical evidence, but still needs dedicated experiments, designed with time-aware behavior in mind).

2. THEORY OF TEMPORAL BEHAVIOR

The speed of interaction is finite [6] both in biological computing and technological computing. The finite speed contributes finite transfering timing between computing units. The effect is well known, but neglected. As pointed out [7], von Neumann’s computing model mentioned the fact that in principle also transfer time must be considered. In the approximation which targeted implementation using vacuum tubes only, it could be neglected. However, the quickly developing technology invalidated that approximation in a stealthy way. For today, timing relations of technological computing are entirely different from those assumed in the classic computing paradigm, which computing science is based upon.

Von Neumann in his famous ‘report’ did provide a “procedure” only for the case when the transmission time can be neglected apart processing time. Actually, von Neumanns statement was that

if [timing relations of] vacuum_tubes
then Classic_Paradigm;
else Unsound;

The more general computing paradigm shall provide the proper procedure instead of the ‘unsound’ branch, when this omission corresponding to [the time relationships between] electron tubes is not valid.

2.1 The General “Procedure”

In the general computing procedure we need to work – in addition to the spatial coordinates – also with the time coordinate of the computing events; considering that those coordinates are connected by the interaction speed. That is, we introduce a 4-D coordinate system, which we call *time-space* system. This representation is in close resemblance with Minkowski’s famous *space-time* system; with the essential difference that a different scale factor is used. In the space-time system all coordinates have dimension of distance, and the time coordinate is given as the corresponding time multiplied by the interaction speed; free and homogeneous propagation is assumed. In our time-space system [8, 6] the spatial distances are divided by the interaction speed and the time represents itself.

That is, all coordinates have dimension of time; this is the appropriate (measurable) quantity given that both technological and biological computing must meet their timing constraints. We assume a sectioned and directed propagation; furthermore the distance is measured along signal path, rather than calculated from the coordinates of its endpoints.

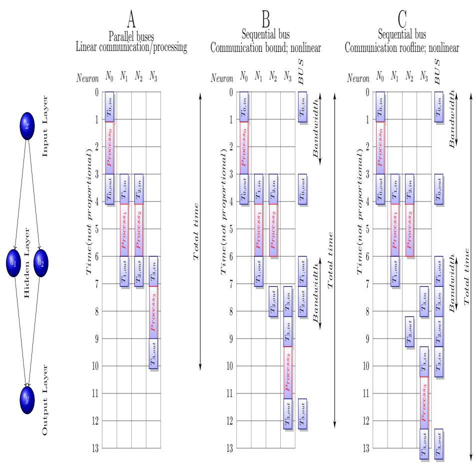

Figure 1: The constrained computing operations in our *time-space* approach. The operating units can be processors, networked computers, gates or neurons; depending on the context. The finite interaction speed and the physical distance of computing elements results in "idle waiting times" (see mixed-color vectors in figure). Idle waiting is one of the major reasons of computing systems' inefficiency.

We describe computing events in this time-space system as constrained vectors, see FIGURE 1. A processing event happens at a given position, but its beginning and end have different time coordinates; that is the processing vector is parallel with axis t. In contrast, a data transfer event changes both spatial and time coordinates simultaneously, that is, the data transfer vectors are not parallel with axis t and are not in the plane perpendicular to it. The vectors are constrained both in the sense that the interaction speed connects their spatial and time coordinates, and that a processing cannot start until its operand arrives, and similarly, delivery of its result cannot start until the processing finished: *the two operations block each other*. This constraint introduces 'idle time' to computing. This idle time in one form leads to inefficiency of processors [9, 7], in another form to inefficiency of parallelized sequential computations [5]. When there are several operands, before the processing may begin, all operands must arrive at their destination.; similarly, all results from multiple processors must be delivered from the output sections of the processing units. These constraints must be emphasized when discussing operation of computing accelerators, especially vector and tensor processors [10, 5].

Biology uses three-state logic, which enables a very low energy consumption. Neurons are "off" until some input is received by their synapses, then for some time they get "open", after some time "inactivated", and again after some time "off" again. As discussed in [17], the three-state operation mode is desirable from several points of view; among others it defines the direction of time. The present design point of view, replacing "inactivated" state with a "down" edge, enables

Table 1: Summary of graphic tablets and software solutions used to acquire the data

Software	Graph Logger		Ductus	
Population	Schools	Hospital	Schools	Hospital
Tablet	Wacom Pro M	Wacom Pro L	Wacom 4M	Wacom 3L
Size (mm)	224×148	311×216	223×139	300×200
Numberof children	258	29	192	101

reaching higher operating frequency [4], but large designs shed light on the limitations, caused by attempting to replace the energetically needed three-state operation [39, 40] with a simplified two-state operating mode.

3. CONCLUSION

Machine learning arised from the intention of mimicking biological learning on technical implementations, having several million times higher transfer and processing speed. However, the speed difference sheds light to a common fallacy of their operation: their design does not consider that chained operations must consider also transfer time; furthermore that the computing operation and the transfer operation are blocking each other. Methods of learning and machine learning have mostly orthonal methods, because of their "implementation": biological operation natively considers transfer time, technical implementation assumes immediate interaction (neglects transfer time).

References

[1] Sterling P, Laughlin S. Principles of Neural Design, 1st ed. The MIT Press;2017.

[2] Floridi L. Information - A Very Short Introduction. Oxford University Press;2010.

[3] Stone JV. Principles of Neural Information Theory. Sebtel Press;2018.

[4] Vegh J. In The 2021 International Conference on Computational Science and Computational Intelligence; Foundations of Computer Science FCS'21. IEEE. 2021;FCS4404, in print.

[5] Vegh J. Revising the Classic Computing Paradigm and Its Technological Implementations. Trans Computers, 2021;14:1-13.

[6] Vegh J. Why do we need to Introduce Temporal Behavior in both Modern Science and Modern Computing, With an Outlook to Researching Modern Effects/Materials and Technologies. Global Journal of Computer Science and Technology: Hardware & Computation. 2020;20:13.

[7] J Vegh. In Proceedings of the 2020 International Conference on Computational Science and Computational Intelligence, CSCI'20, 2020. Las Vegas, Nevada, USA. IEEE Computer Society, 2020:CSCI2019.

[8] https://arxiv.org/abs/2006.01128

[9] Hameed R, Qadeer W, Wachs M, Azizi O, Solomatnikov A, et. al. Understanding Sources of Inefficiency in General-Purpose Chips. ACM SIGARCH Computer Architecture News.2010;38:37-47.

[10] http://arxiv.org/abs/2005.08942

[11] J. von Neumann.First Draft of a Report on The EDVAC. IEEE Annals of the History of Computing.1993;15, 27-75.

[12] http://www.imperial.ac.uk/ wl/teachlo-cal/cuscomp/notes/chapter2.pdf

[13] Johnston D, Sin Wu SM. Foundations of Cellular Neurophysiology. Massachusetts Institute of Technology.1995:710.

[14] Koch C. Biophysics of Computation, Oxford University Press, 1999.

[15] Antle MC, Silver R.Circadian Insights into Motivated Behavior. Trends Neurosci. 2015;28,145-151.

[16] Vegh J, Berki A J. In The 2021 International Conference on Computational Science and Computational Intelligence; Foundations of Computer Science FCS'21.IEEE. 2021; FCS4378, in print.

[17] https://www.preprints.org/manuscript/202103.0414/v1

[18] D. Hebb. The Organization of Behavior, Wiley: New York. 1949.1999;50:437.

[19] Lowel S, Singer W. Selection of Intrinsic Horizontal Connections in the Visual Cortex by Correlated Neuronal Activity. Science. 1992;255,209.

[20] https://www.biorxiv.org/content/10.1101/803577v1

[21] Jordan MI. Artificial Intelligence—The Revolution Hasn't Happened Yet. Harvard Data Science Review. 2019.

[22] Hutson M.Core progress in AI has stalled in some fields. Science. 2020;368:927. 146

[23] Losonczy A, Magee J.Integrative Properties of Radial Oblique Dendrites in Hippocampal CA1 Pyramidal Neurons. Neuron. 2006;50, 291-307.

[24] Furber S, Temple S.Neural Systems Engineering. J R Soc Interface. 2007;4:193.

[25] https://arxiv.org/abs/1705.06963

[26] Markovic D, Mizrahi A, Querlioz D, Grollier J. Physics for Neuromorphic Computing. Nature Reviews Physics. 2020;2;499-510.

[27] Kendall JD, Kumar S. The Building Blocks of a Brain-Inspired Computer. Appl Phys Rev. 2020;7:011305.

[28] Waser R. Advanced Electronics Materials and Novel Devices. Nanoelectronics and Information Technology, 3rd Edition.Wiley. 2012.

[29] https://arxiv.org/abs/2001.01266