

Does Computational Neuroscience need new synaptic learning paradigms?

Johanni Brea and Wulfram Gerstner

School of Computer and Communication Sciences and

Brain Mind Institute, School of Life Sciences

Ecole Polytechnique Fédérale de Lausanne, 1015 Lausanne EPFL

July 19, 2016

Abstract *Computational neuroscience is dominated by a few paradigmatic models, but it remains an open question whether the existing modelling frameworks are sufficient to explain observed behavioural phenomena in terms of neural implementation. We take learning and synaptic plasticity as an example and point to open questions, such as one-shot learning and acquiring internal representations of the world for flexible planning.*

Successful paradigms inspire the thinking of researchers and guide scientific research, yet their success may block independent thinking and hinder scientific progress [1]. Influential learning paradigms in computational neuroscience such as the Hopfield model of associative memory [2], the Bienenstock-Cooper-Munro model for receptive field development [3], or Temporal-Difference Learning for reward-based action learning [4] are of that kind. The question arises whether these and related paradigms in machine learning will be sufficient to account for the variety of learning behaviour observed in nature.

Learning paradigms and learning rules

In classic approaches to machine learning and artificial neural networks, learning from data is formalized in three different paradigms: supervised, unsupervised and reinforcement learning [5, 6, 7, 8]. In supervised learning, each sample data point (e.g., a pixel image or measurements for multiple sensors) comes with a label such as 'this image is a cat', 'this image is a dog' (classification task) or for this configuration of sensory data the correct output is 5.8 (regression task). The objective of supervised learning is to optimize parameters of a machine or mathematical function that takes a data point as input and predicts the output, i.e. that performs a correct classification or prediction. Machine learning has developed powerful models and methods, such as support vector machines [9], Gaussian Processes [10], or

stochastic gradient descent in deep neural networks [11] that allow to minimize the classification or regression error.

In contrast with the above, in unsupervised learning we just have multiple sample data points (pixel images or sensor readings), but no notion of correct or incorrect classification. The typical task of such machine learning algorithms consists of finding a representation of the data that would serve as a useful starting point for further processing. Typical objective functions include compression of the data into a low-dimensional space while maximizing the variance or independence of the data under some normalization constraints. The fields of signal processing and machine learning have developed algorithms such as principal component analysis (PCA) [5], projection pursuit [12], independent component analysis (ICA) [13, 14] and sparse coding [15], that optimize these objective functions.

In reinforcement learning, data is not given, but collected by an agent which receives sparse rewards for some state-action pairs [8]. Temporal-difference (TD) learning methods [16] such as Q-learning [17] and SARSA [18], but also policy gradient methods [19, 20] are the best-studied methods that enable the agent to choose actions that eventually maximize the reward.

In contrast to these purely algorithmic methods of machine learning, any learning method in computational neuroscience should ideally provide a link to the brain. In the neurosciences it is widely accepted that learning observed in humans or animals at the behavioural level corresponds, at the level of biological neural networks, to changes in the synaptic connections between neurons [21, 22].

Classical stimulation protocols for long-term potentiation (LTP) [23, 24, 25], long-term depression (LTD) [26, 27], or spike-timing dependent plasticity [28, 29, 30], inspired by Hebbian learning [31], combine the activation of a presynaptic neuron (or presynaptic pathway) with an activation, depolarization, or chemical manipulation of the postsynaptic neurons, to induce synaptic changes. Numerous synaptic plasticity rules have been developed that are inspired by these experimental data [32, 3, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44]. Generically, in plasticity rules of computational neuroscience the change of a synapse from a neuron j to a neuron i is described as

$$\frac{d}{dt}w_{ij} = F(w_{ij}; s_i, a_j) \quad (1)$$

where w_{ij} is the momentary 'weight' of a synapse, s_i describes the state of the postsynaptic neuron (e.g., its membrane potential, calcium concentration, spike times, or firing rate) and a_j is the activity of the presynaptic neuron [45, 46, 47].

Local plasticity rules of the form (1) can be used to implement a large fraction [44] of known unsupervised learning methods such as PCA [48], ICA [49], Projection pursuit [50], or map formation [33, 51, 52, 5, 36, 37] as well as simple forms of supervised learning, where every neuron receives a direct teaching signal [53, 54, 55]. However, a convincing hypothesis for biologically plausible supervised learning in recurrent or multilayer (deep) spiking neural networks has yet to be proposed (but

see [56, 57, 58, 59, 60, 61]).

A link to reinforcement learning can be established by a slight modification of the Hebbian rule in Eq. (1). Let us suppose that the co-activation of pre- and post-synaptic neurons leaves a slowly (with time constant τ_e) decaying trace e_{ij} at the synapses

$$\frac{d}{dt}e_{ij} = F(w_{ij}; s_i, a_j) - \frac{e_{ij}}{\tau_e} \quad (2)$$

which is transformed into a permanent weight change only if a modulatory signal $M(t)$ confirms the change

$$\frac{d}{dt}w_{ij} = e_{ij}(t) M(t). \quad (3)$$

The two-step learning process described in Eqs. (2) and (3) is consistent with experimental data of synaptic plasticity under the influence of neuromodulators [62, 63, 64, 65, 66] as well as with the concepts of synaptic tagging, capture, and consolidation [67, 68, 69]. Interestingly, most, if not all, of the reinforcement learning algorithms in the class of TD-learning and in the class of policy gradient rules can be cast in the form of Eqs. (2) and (3) [70, 71, 72, 53, 73, 74, 75, 76, 77, 78]. An excellent candidate for the modulating factor M in Eq. (3) is the neuromodulator dopamine, since its activity is correlated with reward signals [4, 79].

Associative memory models [80, 81, 2, 82, 83] have been one of the most influential paradigms of learning and memory in computational neuroscience and inspired numerous theoretical studies, e.g., [84, 85, 86, 87, 88, 89, 90]. Their classification in terms of supervised, unsupervised, or reward-based learning is not straightforward. The reason is that in all the cited studies, learning is supposed to have happened somewhere in the past, while the retrieval of previously learned memories is studied under the assumption of *fixed* synaptic weights. Thus, implicitly this paradigm suggests a modulating factor, similar to M in Eq. (3) that determines whether learning is switched off (for retrieval of existing memories) or on (in the case of novel patterns that need to be learned) [91, 92, 93, 78]. If such a novelty-related modulating factor is missing, the creation of new memories with Hebbian learning rules is difficult [94, 95, 96, 97, 98]. Novelty-related factors combined with a Hebb-like STDP rule have also been studied in models of autoencoders or sequence generators with spiking neurons [99, 100].

The existing paradigms in computational neuroscience continue to trigger interesting research that relates synaptic plasticity to learning behaviour. For example, plasticity rules of the form (1) explain the formation of receptive fields in early sensory processing stages like V1 [101, 44]. Models with modulated Hebbian plasticity as in Eqs. (2) and (3) can explain habitual learning as observed for example in the Morris water maze task [77, 78]. And associative memory models explain some behaviour that depend on episodic memory [98, 102].

Limits of learning rules in computational neuroscience

With the standard paradigms of learning in computational neuroscience reviewed

above in mind, we return to the question of whether these paradigms are sufficient to account for the variety of observed learning behaviour, in particular, one-shot learning and updating acquired representations of the world.

Let us consider the following example. When we hear about a traffic jam on the route from home to work, we can easily adapt our behaviour and take an alternative route. Knowing the cause of the traffic jam, e.g. a road construction site, allows us to decide hours later which route to choose on the way back. In this example, the internal representation consists, first, of possible routes between home and work, second, the position and the cause of the traffic jam, and third, cause-dependent expectations about the duration of traffic jams, e.g. a few hours in case of a small accident, at least a day for a road construction site. These three pieces of information are typically acquired at different moments in life and, presumably, all cause lasting synaptic changes that affect behaviour. Importantly, some events are experienced only once, e.g. the news about the traffic jam, but are sufficient to cause long-lasting memories ('one-shot learning' or 'one-shot memorization').

One view on the traffic jam example is that it requires episodic memory that links the 'what, where and when' of specific events. Many models of episodic memory rely on recurrently connected neural networks that implement an associative memory [102, 103, 104] where specific input cues (e.g., position of an object or event) recall certain object representations. The association of 'what' (e.g. traffic jam caused by a road construction site) with 'where' could be learned by strengthening the connections between the corresponding neurons by up-regulation of 'Hebbian' plasticity under neuromodulation. A temporal ordering (when) of what-where associations could be learned by strengthening connections between subsequently active neurons [86, 102]. In these recurrent neural networks, 'one-shot memorization' has been studied in models of palimpsest memory [105, 106, 107, 108, 109, 110, 111], where the last few patterns in a continuous stream of patterns can be recalled and no catastrophic forgetting is observed.

Such models give a conceptual account for the recall of what-where-when associations given a cue. But are they sufficient to explain the behaviour in the traffic jam example? Maybe partially. Experiencing different types of traffic jams, travelling different routes from home to work, the news about the traffic jam: all these experiences could form 'what, where and when' associations. But key questions remain. How does our brain generate internal cues to recall all relevant information about the specific traffic jam, the possible routes and the typical durations? How does it combine the recalled patterns to decide which route to take? Without an answer to these questions it seems that models of associative memory explain only half of a behaviour that requires episodic memory.

An alternative view on the traffic jam example relies on an acquired representation of space. With unsupervised learning in form of competitive Hebbian synaptic plasticity, navigating agents can learn the receptive fields of place cells [70, 112, 113], such that these cells fire exclusively when the agent is at certain positions [114].

Given these place cells, TD-learning allows to learn position-dependent optimal actions to reach a goal [71, 70, 112]. In these models, the learning time to find the optimal actions is comparable to behavioural learning times, if the agent explores a novel and stationary environment (e.g. the standard reference memory watermaze task [71]). But if a well known environment changes abruptly, as in the traffic jam example, learning in these models is much slower than behavioural learning. In order to match the behavioural learning times, the agent needs to acquire a map of the environment that adds metric or topological information to the internal representation and allows planning (see e.g. the delayed-matching-to-place task in [71]).

Learning a map of the environment is just one example of acquiring domain-specific structure to quickly learn novel tasks. Many more examples exist. People that know to read and write can learn from a single presentation of an unseen character to correctly classify and generate new examples [115]. Having learned the rules of grammar or the hierarchical organization of biological species, people can easily generalize from sparse data, like forming the plural of a novel word or inferring from the fact that 'jays are birds' that 'jays are animals' and that 'jays are not mammals'.

Acquiring internal representations that incorporate such domain-specific structures is possible with abstract algorithmic models in machine learning and artificial intelligence, like model-based reinforcement learning [8, 116] hierarchical Bayesian methods [115, 117] or inductive logic programming [118]. It is, in general, not straightforward to translate these models into neural implementations, but for the specific case of learning maps of the environment, there are interesting propositions [119, 120, 71, 121, 122, 123, 124, 125] that could serve to learn the different routes in the traffic jam example and potentially also the expectations about durations of traffic jams, e.g. with models inspired by dynamic programming [125].

We as computational neuroscientists should aim for an explanation of one-shot learning or the acquisition of internal representations that are tightly constrained by both behavioural and physiological data. Currently it seems out of reach to obtain suitable physiological data from humans. But impressive learning behaviour is also observed in food-storing animals [126, 127, 128, 129]. Westerns scrub-jays encounter a problem very similar to the one in the traffic-jam example: they hide different types of food at different places in their environment, and update their search behaviour based on their expectations about the perishability rates of the different types of food [130]. Furthermore, they were observed to be rule learners in simple matching and oddity tasks [131], they use transitive inference to predict social dominance [132] and re-cache hidden food to prevent pilfering, by remembering which individual watched them during particular caching events [133].

In summary, one-shot learning and the acquisition of internal representations for flexible planning do not yet seem to be satisfactorily explained by the dominant paradigms of learning in computational neuroscience. To make progress in our understanding of such flexible learning behaviour, abstract models on an algorithmic

level could give hints for novel models of synaptic learning that then, in turn, need to be constrained by physiological and behavioural data.

Acknowledgments

We thank Alexander Seeholzer for a careful reading of the manuscript and constructive comments. Research was supported by the European Research Council grant no. 268689 (MultiRules).

References

- [1] T. S. Kuhn. *The structure of scientific revolutions*. University of Chicago Press, 1970.
- [2] J. J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. USA*, 79:2554–2558, 1982.
**** A classic and highly influential paper on attractor networks for associative memory.**
- [3] E. L. Bienenstock, L. N. Cooper, and P. W. Munroe. Theory of the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. *J. Neurosci.*, 2:32–48, 1982.
**** A classic paper for explaining receptive field development with a synaptic plasticity (BCM) rule. Still many modern plasticity rules with spiking neurons relate to the BCM rule [41, 42, 43].**
- [4] W. Schultz, P. Dayan, and R.R. Montague. A neural substrate for prediction and reward. *Science*, 275:1593–1599, 1997.
**** A classic paper linking the theory of reinforcement learning to learning behaviour and neural activity by interpreting experimentally found dopaminergic activity as an implementation of the reward prediction error of temporal difference learning.**
- [5] J. Hertz, A. Krogh, and R. G. Palmer. *Introduction to the Theory of Neural Computation*. Addison-Wesley, Redwood City CA, 1991.
- [6] S. Haykin. *Neural Networks*. Prentice Hall, Upper Saddle River, NJ, 1994.
- [7] C. M. Bishop. *Neural Networks for Pattern Recognition*. Clarendon Press, Oxford, 1995.
- [8] R. Sutton and A. Barto. *Reinforcement learning*. MIT Press, Cambridge, 1998.
- [9] B. Schölkopf and A. Smola. *Learning with kernels: support vector machines, regularization, optimization, and beyond*. MIT Press Cambridge, 2002.

- [10] C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*. The MIT Press, 2006.
- [11] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 25*, pages 1097–1105. Curran Associates, Inc., 2012.
- [12] J. Friedman. Exploratory projection pursuit. *J. American Statistical Assoc.*, 82:249–266, 1987.
- [13] C. Bell, V. Han, Y. Sugawara, and K. Grant. Synaptic plasticity in a cerebellum-like structure depends on temporal order. *Nature*, 387:278–281, 1997.
- [14] A. Hyvärinen, J. Karhunen, and E. Oja. *Independent Component Analysis*. Wiley, 2001.
- [15] B. A. Olshausen and D. J. Field. Natural image statistics and efficient coding. *Network: Computation in Neural Systems*, 7:333–339, 1996.
- [16] P. Dayan. The convergence of TD(λ) for general λ . *Machine learning*, 8:341–362, 1992.
- [17] C. J. C. H. Watkins and P. Dayan. Q-learning. *Machine Learning*, 8:279–292, 1992.
- [18] G. A. Rummery and M. Niranjan. *Online Q-learning using connectionist systems*. Cambridge University, 1994.
- [19] R. Williams. Simple statistical gradient-following methods for connectionist reinforcement learning. *Machine Learning*, 8:229–256, 1992.
- [20] J. Baxter, P. Bartlett, and L. Weaver. Experiments with infinite-horizon, policy-gradient estimation. *Journal of Artificial Intelligence Research*, 15:351–381, 2001.
- [21] R. G. M. Morris, E. Anderson, G. S. Lynch, and M. Baudry. Selective impairment of learning and blockade of long-term potentiation by an n-methyl-d-aspartate receptor antagonist, ap5. *Nature*, 319:774–776, 1986.
- [22] S. Martin, P. Grimwood, and R. Morris. Synaptic plasticity and memory: an evaluation of the hypothesis. *Ann. Rev. Neurosci.*, 23:649–711, 2000.
- [23] T. Bliss and A. Gardner-Medwin. Long-lasting potentiation of synaptic transmission in the dentate area of unanaesthetized rabbit following stimulation of the perforant path. *J. Physiol.*, 232:357–374, 1973.
- [24] R. C. Malenka and R. A. Nicoll. Long-term potentiation—a decade of progress? *Science*, 285:1870–1874, 1999.
- [25] J. Lisman. Long-term potentiation: outstanding questions and attempted synthesis. *Phil. Trans. R. Soc. Lond B: Biological Sciences*, 358:829 – 842, 2003.
- [26] G. Lynch, T. Dunwiddie, and V. Gribkoff. Heterosynaptic depression: a postsynaptic correlate of long-term potentiation. *Nature*, 266:737–739, 1977.

- [27] W. B. Levy and D. Stewart. Temporal contiguity requirements for long-term associative potentiation/depression in hippocampus. *Neurosci.*, 8:791–797, 1983.
- [28] H. Markram, J. Lübke, M. Frotscher, and B. Sakmann. Regulation of synaptic efficacy by coincidence of postsynaptic AP and EPSP. *Science*, 275:213–215, 1997.
- [29] G. Bi and M. Poo. Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. *J. Neurosci.*, 18:10464–10472, 1998.
- [30] P. Sjöström, G. Turrigiano, and S. Nelson. Rate, timing, and cooperativity jointly determine cortical synaptic plasticity. *Neuron*, 32:1149–1164, 2001.
- [31] D. O. Hebb. *The Organization of Behavior*. Wiley, New York, 1949.
- [32] D. J. Willshaw and C. von der Malsburg. How patterned neuronal connections can be set up by self-organization. *Proc. R. Soc. (London) Ser. B*, 194:431–445, 1976.
- [33] T. Kohonen. *Self-organization and associative memory, 3rd edition*. Springer-Verlag, Berlin Heidelberg New York, 1989.
- [34] R. Linsker. From basic network principles to neural architecture: emergence of spatial-opponent cells. *Proc. Natl. Acad. Sci. USA*, 83:7508–7512, 1986.
- [35] K. Miller, J. B. Keller, and M. P. Stryker. Ocular dominance column development: analysis and simulation. *Science*, 245:605–615, 1989.
- [36] E. Erwin, K. Obermayer, and K. Schulten. Models of orientation and ocular dominance columns in the visual cortex: a critical comparison. *Neural Comput.*, 7:425–468, 1995.
- [37] L. Wiskott and T. Sejnowski. Constraint optimization for neural map formation: a unifying framework for weight growth and normalization. *Neural Computation*, 10:671–716, 1998.
- [38] R. Kempter, W. Gerstner, and J. L. van Hemmen. Hebbian learning and spiking neurons. *Phys. Rev. E*, 59:4498–4514, 1999.
- [39] S. Song, K. Miller, and L. Abbott. Competitive Hebbian learning through spike-time-dependent synaptic plasticity. *Nature Neuroscience*, 3:919–926, 2000.
- [40] W. Senn, H. Markram, and M. Tsodyks. An algorithm for modifying neurotransmitter release probability based on pre- and postsynaptic spike timing. *Neural Computation*, 13:35–67, 2000.
- [41] H. Z. Shouval, M. F. Bear, and L. N. Cooper. A unified model of NMDA receptor-dependent bidirectional synaptic plasticity. *Proceedings of the National Academy of Sciences of the United States of America*, 99:10831–6, 2002.
- [42] J.-P. Pfister and W. Gerstner. Triplets of spikes in a model of spike timing-dependent plasticity. *J. Neuroscience*, 26:9673–9682, 2006.

- [43] C. Clopath, L. Busing, E. Vasilaki, and W. Gerstner. Connectivity reflects coding: A model of voltage-based spike-timing-dependent-plasticity with homeostasis. *Nature Neuroscience*, 13:344–352, 2010.
- [44] C. S. Brito. Nonlinear hebbian learning as a unifying principle in receptive field formation. *PLOS Comput. Biol.*, xx:xx, 2016.
- * Shows that plasticity models motivated by sparse coding, ICA and projection pursuit have a natural interpretation as non-linear Hebbian learning, which explains receptive field development, independent of the details of the non-linearity.**
- [45] T. H. Brown, A. M. Zador, Z. F. Mainen, and B. J. Claiborne. Hebbian modifications in hippocampal neurons. In M. Baudry and J. Davis, editors, *Long-term potentiation.*, pages 357–389. MIT Press, Cambridge, London, 1991.
- [46] A. Morrison, M. Diesmann, and W. Gerstner. Phenomenological models of synaptic plasticity based on spike timing. *Biol. Cybern.*, 98:459–478, 2008.
- [47] W. Gerstner, W. Kistler, R. Naud, and L. Paninski. *Neuronal Dynamics. From single neurons to networks and cognition.* Cambridge Univ. Press, 2014.
- [48] E. Oja. A simplified neuron model as a principal component analyzer. *J. Mathematical Biology*, 15:267–273, 1982.
- [49] A. Hyvarinen and E. Oja. Independent component analysis by general nonlinear hebbian-like learning rules. *Signal Processing*, 64:301–313, 1998.
- [50] C. Fyfe and R. Baddeley. Non-linear data structure extraction using simple hebbian networks. *Biol. Cybern.*, 72:533–541, 1995.
- [51] T. Kohonen. Physiological interpretation of the self-organizing map algorithm. *Neural Networks*, 6:895–905, 1993.
- [52] G. Carpenter. Distributed learning, recognition and prediction by ART and ARTMAP neural networks. *Neural Networks*, 10:1473–1494, 1997.
- [53] J.-P. Pfister, T. Toyoizumi, D. Barber, and W. Gerstner. Optimal spike-timing dependent plasticity for precise action potential firing in supervised learning. *Neural Computation*, 18:1318–1348, 2006.
- [54] R. Urbanczik and W. Senn. Learning by the Dendritic Prediction of Somatic Spiking. *Neuron*, 81:521–528, 2014.
- [55] R. Güttig. Spiking neurons can discover predictive features by aggregate-label learning. *Science*, 351:aab4113–1, 2016.
- [56] K. D. Harris. Stability of the fittest: organizing learning through retroaxonal signals. *Trends in Neurosciences*, 31:130–136, 2008.
- [57] D. Sussillo and L. F. Abbott. Generating coherent patterns of activity from chaotic neural networks. *Neuron*, 63:544–57, 2009.

- [58] G. M. Hoerzer, R. Legenstein, and W. Maass. Emergence of complex computational structures from chaotic neural networks through reward-modulated hebbian learning. *Cerebral Cortex*, 24:677–690, 2014.
- [59] M. Schiess, R. Urbanczik, and W. Senn. Somato-dendritic Synaptic Plasticity and Error-backpropagation in Active Dendrites. *PLoS Comput Biol*, 12:e1004638, 2016.
- [60] Y. Bengio, D.-H. Lee, J. Bornschein, and Z. Lin. Towards Biologically Plausible Deep Learning. *arXiv:1502.0415*, 2015.
- [61] B. Scellier and Y. Bengio. Towards a Biologically Plausible Backprop. *arXiv:1602.05179*, 2016.
- [62] J. Reynolds and J. Wickens. Dopamine-dependent plasticity of corticostriatal synapses. *Neural Networks*, 15:507–521, 2002.
- [63] G. Seol, J. Ziburkus, S. Huang, L. Song, I. Kim, K. Takamiya, R. Huganir, H.-K. Lee, and A. Kirkwood. Neuromodulators control the polarity of spike-timing-dependent synaptic plasticity. *Neuron*, 55:919–929, 2007.
- [64] J. Zhang, P. Lau, and G. Bi. Gain in sensitivity and loss in temporal contrast of stdp by dopaminergic modulation at hippocampal synapses. *Proc. Natl. Aca. Sci. USA*, 106:13028–13033, 2009.
- [65] V. Pawlak and J. Kerr. Dopamine receptor activation is required for corticostriatal spike-timing-dependent plasticity. *J. Neuroscience*, 28:2435–2446, 2008.
- [66] V. Pawlak, J. Wickens, A. Kirkwood, and J. Kerr. Timing is not everything: neuromodulation opens the STDP gate. *Front. Synaptic Neurosci.*, 2:146, 2010.
- [67] U. Frey and R. Morris. Synaptic tagging and long-term potentiation. *Nature*, 385:533–536, 1997.
- [68] R. L. Redondo and R. G. M. Morris. Making memories last: the synaptic tagging and capture hypothesis. *Nat Rev Neurosci*, 12:17–30, 2011.
- [69] J. Lisman, A. A. Grace and E. Duzel. A neoHebbian framework for episodic memory; role of dopamine-dependent late LTP. *Trends in Neurosciences*, 34:536–547, 2011.
- [70] A. Arleo and W. Gerstner. Spatial cognition and neuro-mimetic navigation: a model of hippocampal place cell activity. *Biological Cybernetics*, 83:287–299, 2000.
- [71] D. Foster, R. Morris, and P. Dayan. Models of hippocampally dependent navigation using the temporal difference learning rule. *Hippocampus*, 10:1–16, 2000.
- [72] X. Xie and S. Seung. Learning in neural networks by reinforcement of irregular spiking. *Phys. Rev. E*, 69:41909, 2004.
- [73] R. V. Florian. Reinforcement learning through modulation of spike-timing-dependent synaptic plasticity. *Neural Computation*, 19:1468–1502, 2007.
- [74] E. Izhikevich. Solving the distal reward problem through linkage of STDP and dopamine signaling. *Cerebral Cortex*, 17:2443–2452, 2007.

- [75] R. Legenstein, D. Pecevski, and W. Maass. Theoretical analysis of learning with reward-modulated spike-timing-dependent plasticity. In J. Platt, D. Koller, Y. Singer, and S. Roweis, editors, *Advances in Neural Information Processing Systems 20*. MIT Press, Cambridge, MA, 2008.
- [76] N. Frémaux, H. Sprekeler, and W. Gerstner. Functional requirements for reward-modulated spike-timing-dependent plasticity. *J. Neurosci.*, 40:13326–13337, 2010.
- [77] N. Frémaux, H. Sprekeler, and W. Gerstner. Reinforcement learning using a continuous time actor-critic framework with spiking neurons. *PLoS Comput. Biol.*, 9:e1003024, 2013.
- [78] N. Frémaux and W. Gerstner. Neuromodulated Spike-Timing-Dependent Plasticity and Theory of Three-Factor Learning Rules. *Frontiers in Neural Circuits*, 9:85, 2016.
*** A review on neuromodulated STDP that explains how such plasticity rules implement reward-based and novelty-based learning.**
- [79] W. Schultz. Getting formal with dopamine and reward. *Neuron*, 36:241–263, 2002.
- [80] D. J. Willshaw, O. P. Bunemann, and H. C. Longuet-Higgins. Non-holographic associative memory. *Nature*, 222:960–962, 1969.
- [81] W. A. Little and G. L. Shaw. Analytical study of the memory storage capacity of a neural network. *Math. Biosci.*, 39:281–290, 1978.
- [82] J. J. Hopfield. Neurons with graded response have computational properties like those of two-state neurons. *Proc. Natl. Acad. Sci. USA*, 81:3088–3092, 1984.
- [83] D. J. Amit, H. Gutfreund, and H. Sompolinsky. Storing infinite number of patterns in a spin-glass model of neural networks. *Phys. Rev. Lett.*, 55:1530–1533, 1985.
- [84] D. J. Amit, H. Gutfreund, and H. Sompolinsky. Information storage in neural networks with low levels of activity. *Phys. Rev. A*, 35:2293–2303, 1987.
- [85] M. Tsodyks and M. Feigelman. The enhanced storage capacity in neural networks with low activity level. *Europhys. Lett.*, 6:101–105, 1988.
- [86] A. V. M. Herz, B. Sulzer, R. Kühn, and J. L. van Hemmen. Hebbian learning reconsidered: Representation of static and dynamic objects in associative neural nets. *Biol. Cybern.*, 60:457–467, 1989.
- [87] D. J. Amit and M. V. Tsodyks. Quantitative study of attractor neural networks retrieving at low spike rates. I: Substrate — spikes, rates, and neuronal gain. *Network*, 2:259–273, 1991.
- [88] W. Gerstner and J. L. van Hemmen. Associative memory in a network of ‘spiking’ neurons. *Network*, 3:139–164, 1992.
- [89] D. J. Amit and N. Brunel. A model of spontaneous activity and local delay activity during delay periods in the cerebral cortex. *Cerebral Cortex*, 7:237–252, 1997.
- [90] N. Brunel. Dynamics of sparsely connected networks of excitatory and inhibitory neurons. *Computational Neuroscience*, 8:183–208, 2000.

- [91] M. Hasselmo. The role of acetylcholine in learning and memory. *Curr. Opinion Neurobiol.*, 16:710–715, 2006.
- [92] Q. Gu. Neuromodulatory transmitter systems in the cortex and their role in cortical plasticity. *Neuroscience*, 111:815–835, 2002.
- [93] D. Moncada and H. Viola. Induction of long-term memory by exposure to novelty requires protein synthesis: Evidence for a behavioral tagging. *J. Neurosci*, 27:7476–7481, 2007.
- [94] S. Fusi. Hebbian spike-driven synaptic plasticity for learning patterns of mean firing rates. *Biol. Cybern.*, 87:459–470, 2002.
- [95] G. Mongillo, E. Curti, S. Romani, and D. Amit. Learning in realistic networks of spiking neurons and spike-driven plastic synapses. *Europ. J. Neurosci.*, 21:3143–3160, 2005.
- [96] S. Fusi and L. Abbott. Limits on the memory storage capacity of bounded synapses. *Nature Neuroscience*, 10:485–493, 2007.
- [97] A. Litwin-Kumar and B. Doiron. Formation and maintenance of neuronal assemblies through synaptic plasticity. *Nature Commun.*, 5:5319, 2014.
- [98] F. Zenke, E. Agnes, and W. Gerstner. Diverse synaptic plasticity mechanisms orchestrated to form and retrieve memories in spiking neural networks. *Nature Commun.*, 6:6922, 2015.
- * A model of associative memory with spiking neurons and with a synaptic plasticity mechanism that is robust against ongoing activity.**
- [99] J. Brea, W. Senn, and J.-P. Pfister. Matching recall and storage in sequence learning with spiking neural networks. *J. Neuroscience*, 33:9565–9575, 2013.
- [100] D. Rezende and W. Gerstner. Stochastic variational learning in recurrent spiking networks. *Frontiers in Computational Neuroscience*, 8:38, 2014.
- [101] H. Ko, L. Cossell, C. Baragli, J. Antolik, C. Clopath, S. B. Hofer, and T. D. Mrsic-Flogel. The emergence of functional microcircuits in visual cortex. *Nature*, 496:96–100, 2013.
- * Reports the emergence of microcircuits in the visual cortex of mice after eye opening and explains this with an activity dependent synaptic plasticity rule.**
- [102] R. P. Kesner and E. T. Rolls. A computational theory of hippocampal function, and tests of the theory: New developments. *Neuroscience and Biobehavioral Reviews*, 48:92–147, 2015.
- * Detailed review of a model of episodic memory based on an attractor network.**
- [103] K. Norman, D. Greg, and S. M. Polyn. Computational Models of Episodic Memory. In *The Cambridge Handbook of Computational Cognitive Modeling*, pages 189–224. Cambridge University Press, 2008.

- [104] E. T. Rolls, S. M. Stringer, and T. P. Trappenberg. A unified model of spatial and episodic memory. *Proceedings of the Royal Society B: Biological Sciences*, 269:1087–1093, 2002.
- [105] J. Nadal, G. Toulouse, J. Changeux, and S. Dehaene. Networks of Formal Neurons and Memory Palimpsests. *Europhys. Lett.*, 1:535–542, 1986.
- [106] D. J. Amit and S. Fusi. Learning in Neural Networks with Material Synapses. *Neural computation*, 6:957–982, 1994.
- [107] S. Fusi, P. J. Drew, and L. F. Abbott. Cascade models of synaptically stored memories. *Neuron*, 45:599–611, 2005.
- [108] W. Senn and S. Fusi. Convergence of stochastic learning in perceptrons with binary synapses. *Phys, Rev. E*, 71:061907, 2005.
- [109] M. Papper, R. Kempter, and C. Leibold. Synaptic tagging, evaluation of memories, and the distal reward problem. *Learn. Mem.*, 18:58–70, 2011.
- [110] Y. Amit and Y. Huang. Precise Capacity Analysis in Binary Networks with Multiple Coding Level Inputs. *Neural Comput*, 22:660–688, 2010.
- [111] T. Elliott. Discrete States of Synaptic Strength in a Stochastic Model of Spike-Timing-Dependent Plasticity. *Neural Comput*, 22:244–272, 2010.
- [112] D. Sheynikhovich, R. Chavarriaga, T. Strosslin, A. Arleo, and W. Gerstner. Is there a geometric module for spatial orientation? insights from a rodent navigation model. *Psychological Review*, 116:540–566, 2009.
- [113] M. Franzius, H. Sprekeler and L. Wiskott. Slowness and sparseness lead to place, head-direction, and spatial-view cells. *PLoS Comput. Biol.* 3:e166, 2007.
- [114] J. O’Keefe and J. Dostrovsky. The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain research*, 34:171–5, 1971.
- [115] B. M. Lake, R. Salakhutdinov, and J. B. Tenenbaum. Human-level concept learning through probabilistic program induction. *Science*, 350:1332–1338, 2015.
- * Reports one-shot learning of handwritten characters by humans.**
- [116] A. Guez, D. Silver, and P. Dayan. Scalable and efficient bayes-adaptive reinforcement learning based on Monte-Carlo tree search. *Journal of Artificial Intelligence Research*, 48:841–883, 2013.
- [117] C. Kemp and J. B. Tenenbaum. The discovery of structural form. *Proceedings of the National Academy of Sciences of the United States of America*, 105:10687–92, 2008.
- * A Bayesian approach to discover structure and form in data that shows similarities with how scientists and children discover forms like hierarchies, cliques or relations.**
- [118] S. Muggleton and L. de Raedt. Inductive logic programming: theory and methods. *Journal of Logic Programming*, 19:629–679, 1994.

- [119] K. Blum and L. Abbott. A model of spatial map formation in the hippocampus of the rat. *Neural Comput.*, 8:85–93, 1996.
- [120] W. Gerstner and L. F. Abbott. Learning navigational maps through potentiation and modulation of hippocampal place cells. *J. Comput. Neurosci.*, 4:79–94, 1997.
- [121] K. L. Stachenfeld, M. M. Botvinick, and S. J. Gershman. Design Principles of the Hippocampal Cognitive Map. *Advances in Neural Information Processing Systems 27*, pages 1–9, 2014.
- [122] D. S. Corneil and W. Gerstner. Attractor Network Dynamics Enable Preplay and Rapid Path Planning in Mazelike Environments. *Advances in Neural Information Processing Systems 28*, pages 1675–1683, 2015.
- [123] M. J. Milford, G. F. Wyeth, and D. Prasser. RatSLAM: A Hippocampal Model for Simultaneous Localization and Mapping. *Proceeding of the 2004 IEEE international Conference on Robotics & Automation*, pages 403–408, 2004.
- [124] W. D. Penny, P. Zeidman, and N. Burgess. Forward and Backward Inference in Spatial Cognition. *PLoS Computational Biology*, 9:e1003383, 2013.
- [125] J. Friedrich and M. Lengyel. Goal-Directed Decision Making with Spiking Neurons. *Journal of Neuroscience*, 36:1529–1546, 2016.
- * A spiking neural network model of learning a representation of the world and using it for planning.**
- [126] S. B. Vander Wall. Food Hoarding in Animals. *Chicago: University of Chicago Press.*, 1990.
- [127] D. W. Macdonald. Food Caching by Red Foxes and Some Other Carnivores. *Zeitschrift für Tierpsychologie*, 42:170–185, 1976.
- [128] N. S. Clayton and J. R. Krebs. Memory for spatial and object-specific cues in food-storing and non-storing birds. *Journal of Comparative Physiology A*, 174:371–379, 1994.
- [129] N. S. Clayton and N. J. Emery. Avian Models for Human Cognitive Neuroscience: A Proposal. *Neuron*, 86:1330–1342, 2015.
- ** reviews the fascinating learning abilities of corvids and other birds.**
- [130] N. S. Clayton, K. S. Yu, and A. Dickinson. Interacting Cache memories: evidence for flexible memory use by Western Scrub-Jays (*Aphelocoma californica*). *Journal of experimental psychology. Animal behavior processes*, 29:14–22, 2003.
- [131] B. Wilson, N. J. Mackintosh, and R. A. Boakes. Transfer of relational rules in matching and oddity learning by pigeons and corvids. *The Quarterly Journal of Experimental Psychology Section B*, 37:313–332, 1985.
- [132] C. G. Paz-y Miño, A. B. Bond, A. C. Kamil, and R. P. Balda. Pinyon jays use transitive inference to predict social dominance. *Nature*, 430:778–781, 2004.
- [133] J. M. Dally, N. J. Emery, and N. S. Clayton. Food-caching western scrub-jays keep track of who was watching when. *Science*, 312:1662–1665, 2006.