

cases	doc_1		doc_2		decision	id
			<div>authors<div><ul style="list-style-type: none">Ivana KajiChris EliasmithAaron Voelker</div></div>		DUPLICATES	307
			<div>title</div> Legendre Memory Units: Continuous-Time Representation in Recurrent Neural Networks			
			<div>publication_date</div> 2019-12-01 00:00:00			
			<div>source</div> SupportedSources.PAPERS_WITH_CODE			
			<div>journal</div>			
			<div>volume</div>			
			<div>doi</div>			
			<div>urls</div> <ul style="list-style-type: none">http://papers.nips.cc/paper/9689-legendre-memory-units-continuous-time-representation-in-recurrent-neural-networks.pdfhttps://github.com/abr/neurips2019			
			<div>id</div> id-4391443143788058392			
			<div>abstract</div> <p>We propose a novel memory cell for recurrent neural networks that dynamically maintains information across long windows of time using relatively few resources. The Legendre Memory Unit~(LMU) is mathematically derived to orthogonalize its continuous-time history -- doing so by solving \$d\$ coupled ordinary differential equations~(ODEs), whose phase space linearly maps onto sliding windows of time via the Legendre polynomials up to degree \$d - 1\$. Backpropagation across LMUs outperforms equivalently-sized LSTMs on a chaotic time-series prediction task, improves memory capacity by two orders of magnitude, and significantly reduces training and inference times. LMUs can efficiently handle temporal dependencies spanning \$100\text{,}000\$ time-steps, converge rapidly, and use few internal state-variables to learn complex functions spanning long windows of time -- exceeding state-of-the-art performance among RNNs on permuted sequential MNIST. These results are due to the network's disposition to learn scale-invariant features independently of step size. Backpropagation through the ODE solver allows each layer to adapt its internal time-step, enabling the network to learn task-relevant time-scales. We demonstrate that LMU memory cells can be implemented using \$m\$ recurrently-connected Poisson spiking neurons, \$\mathcal{O}(m)\$ time and memory, with error scaling as \$\mathcal{O}(d / \sqrt{m})\$. We discuss implementations of LMUs on analog and digital neuromorphic hardware.</p>			
			<div>versions</div>			