# Learning Sequential Structure in Simple Recurrent Networks

*Authors: David Servan-Schreiber, Axel Cleeremans, James McClelland*

Abstract: We explore a network architecture introduced by Elman (1988) for predicting successive elements of a sequence. The network uses the pattern of activation over a set of hidden units from time-step t-l, together with element t, to predict element t+ 1. When the network is trained with strings from a particular finite-state grammar, it can learn to be a perfect finite-state recognizer for the grammar. Cluster analyses of the hidden-layer patterns of activation showed that they encode prediction-relevant information about the entire path traversed through the network. We illustrate the phases of learning with cluster analyses performed at different points during training.Several connectionist architectures that are explicitly constrained to capture sequential infonnation have been developed. Examples are Time Delay Networks (e.g. Sejnowski & Rosenberg. 1986) -- also called 'moving window' paradigms -- or algorithms such as back-propagation in time (Rumelhart. Hinton & Williams. 1986), Such architectures use explicit representations of several consecutive events. if not of the entire history of past inputs. Recently. Elman (1988) has introduced a simple recurrent network (SRN) that has the potential to master an infinite corpus of sequences with the limited means of a learning procedure that is completely local in time (see Figure I.).