

Modeling serial recall in working memory with recurrent neural networks

Auteur : Stordeur, Lucas

Promoteur(s) : Sacré, Pierre

Faculté : Faculté des Sciences appliquées

Diplôme : Master en ingénieur civil biomédical, à finalité spécialisée

Année académique : 2024-2025

URI/URL : <http://hdl.handle.net/2268.2/23282>

Avertissement à l'attention des usagers :

Tous les documents placés en accès ouvert sur le site le site MatheO sont protégés par le droit d'auteur. Conformément aux principes énoncés par la "Budapest Open Access Initiative"(BOAI, 2002), l'utilisateur du site peut lire, télécharger, copier, transmettre, imprimer, chercher ou faire un lien vers le texte intégral de ces documents, les disséquer pour les indexer, s'en servir de données pour un logiciel, ou s'en servir à toute autre fin légale (ou prévue par la réglementation relative au droit d'auteur). Toute utilisation du document à des fins commerciales est strictement interdite.

Par ailleurs, l'utilisateur s'engage à respecter les droits moraux de l'auteur, principalement le droit à l'intégrité de l'oeuvre et le droit de paternité et ce dans toute utilisation que l'utilisateur entreprend. Ainsi, à titre d'exemple, lorsqu'il reproduira un document par extrait ou dans son intégralité, l'utilisateur citera de manière complète les sources telles que mentionnées ci-dessus. Toute utilisation non explicitement autorisée ci-avant (telle que par exemple, la modification du document ou son résumé) nécessite l'autorisation préalable et expresse des auteurs ou de leurs ayants droit.

Modeling serial recall in working memory with recurrent neural networks

Stordeur Lucas

University of Liège

Supervisors: Pr. Pierre Sacré, Pr. Steve Majerus

June 10, 2025

Abstract

In psychology, various tasks are designed to study cognitive functions such as language processing, long-term memory, and working memory. Immediate serial recall is one such task, where participants are presented with a sequence of items (e.g., words or letters) and must recall them in the same order after the sequence is presented. This task provides valuable insights into working memory mechanisms.

Computational modeling of psychological experiments helps improve our understanding of working memory by replicating experimental data. Different models exist for serial recall, each based on distinct mechanisms and assumptions that explain various observed behaviors. The quality of these models is usually assessed by comparing their predictions with real experimental data. Among these models, RNNs stand out due to their dynamical properties and strong empirical support from neuroscience. An RNN consists of three main layers: an input layer, where sequence items are presented; a hidden layer, where information is processed; and an output layer, which generates predictions. A key feature of RNNs is their recurrent connections, allowing past inputs to influence future predictions, making them well-suited for capturing temporal relationships in sequences.

The first objective of my thesis is to construct an RNN capable of reproducing three key effects observed in serial recall experiments:

length-dependent performance (recall accuracy decreases sigmoidally as sequence length increases), primacy and recency effects (better recall of the first and last items in a sequence), transposition errors (recall mistakes typically involve swapping items) and to explore the reproduction of many other effects characteristic of this task. To achieve this, I formally modeled the serial recall task and then trained the RNN. Unlike standard deep learning approaches that optimize for maximum accuracy, our model is trained to match the recall performance observed in human participants. This requires two key modifications: (1) training the network only until it reaches human-like recall accuracy and (2) varying the dataset size to prevent overfitting to specific sequences. We will show that despite the difficulties that many encountered to construct such a model, recurrent neural networks are able to reproduce different empirical benchmarks.

The second objective is to explore the internal dynamics of the networks. Specifically, we investigate whether modifying the cell dynamics to be more biologically plausible improves the model alignment with experimental data. The hidden layer activation dynamics—determined by the interaction between current inputs and past hidden states—are progressively refined. I begin with a simple Elman network, then introduce more complex architectures: Gated Recurrent Unit (GRU) cells, which incorporate a forgetting mechanism, Bistable Recurrent Cells (BRCs), which introduce bistability properties. We will show that more complex and biologically plausible allows to achieve better results on the empirical benchmarks.

Finally, as this model has not yet been widely explored, we explored its behavior by investigating how the variability can be included in the curves, as they are a central characteristic of the human data, but is not well explored in model. In addition, the modelization of the inter-patient variability. In addition, saturation of model training will be investigated and will allow to create a more plausible model. All of these contributes to the better understanding of the model.

In conclusion, through these refinements, we aimed to bridge the gap between artificial and biological neural networks, contributing to a deeper understanding of working memory.