# Article information:

Encoding Recurrence into Transformers | OpenReview
<https://openreview.net/forum?id=7YfHla7IxBJ>

# Article summary:

1. This paper proposes a new module, Self-Attention with Recurrence (RSA), to encode the recurrent dynamics of an RNN layer into Transformers.

2. RSA can achieve higher sample efficiency than its corresponding baseline Transformer by leveraging the recurrent inductive bias of Recurrence Encoding Matrices (REMs).

3. A data-driven gated mechanism is used to control the relative proportions of self-attention and REMs in RSA modules, and their effectiveness is demonstrated by four sequential learning tasks.

# Article rating:

May be slightly imbalanced: The article presents the information in a generally reliable way, but there are minor points of consideration that could be explored further or claims that are not fully backed by appropriate evidence. Some perspectives may also be omitted, and you are encouraged to use the research topics section to explore the topic further.

# Article analysis:

The article appears to be trustworthy and reliable overall, as it provides a detailed description of the proposed module and its effectiveness on four sequential learning tasks. The authors also provide evidence for their claims, such as demonstrating how RSA can achieve higher sample efficiency than its corresponding baseline Transformer. Furthermore, the article does not appear to contain any promotional content or partiality towards any particular point of view.

However, there are some points that could be further explored in order to make the article more comprehensive. For example, while the authors mention that RSA can achieve higher sample efficiency than its corresponding baseline Transformer, they do not provide any evidence for this claim or discuss possible risks associated with using RSA instead of a baseline Transformer. Additionally, while the authors discuss how a data-driven gated mechanism is used to control the relative proportions of self-attention and REMs in RSA modules, they do not explore other potential methods for controlling these proportions or discuss why this method was chosen over others. Finally, while the authors demonstrate how effective RSA modules are on four sequential learning tasks, they do not explore other potential applications for this module or discuss how it could be further improved upon in future research.

# Topics for further research:

* Risk assessment of RSA modules
* Alternative methods for controlling self-attention and REMs
* Potential applications of RSA modules
* Improving RSA modules
* Sample efficiency of RSA modules
* Data-driven gated mechanism for RSA modules

# Report location:

<https://www.fullpicture.app/item/fc790d358acfa521a774c8e96a224421>