Mamba & Samba

Title: Comparison of Mamba and Samba in Transformer Models

Subtitle: "Efficient Sequence Processing with SSM and Hybrid Models"

By ChatGPT4.0o, 2025-02-13

Artificial Intelligence

Creating the Future

Dong-A University

Division of Computer Engineering & Artificial Intelligence

Introduction

What are Mamba and Samba?

- **Mamba**: A **State Space Model (SSM)**-based architecture that **removes self-attention** for **efficient sequence processing.**

- Paper : Linear-Time Sequence Modeling with Selective State Spaces

- **Samba**: A **hybrid model** combining **Mamba**'s efficiency with **Transformer's self-attentio**n for balanced performance.

- Paper : Simple Hybrid State Space Models for Efficient Unlimited Context Language Modeling

Mamba & Samba – Key Features

➢ What is Mamba?

- A transformer alternative that **removes self-attention**
- Uses **State Space Models (SSM)** for efficient long-sequence processing

☑ Advantages:

- **O(N) complexity** (better than **Transformer's O(N^2)**)
- **Lower memory** usage & **faster computation**
- Excels in handling **long sequences** (text, audio, time-series)

☒ Disadvantages:

- May **underperform in some NLP tasks** compared to Transformers
- **Lacks direct self-attention capabilities**

➢ What is Samba?

- A **hybrid model** that combines **Mamba's SSM** with **Transformer's self-attention**
- Uses **self-attention for short sequences** and **SSM for long sequences**

☑ Advantages:

- **Maintains Transformer's expressive power**
- Works well for **both short & long sequences**
- More **balanced performance** across tasks

☒ Disadvantages:

- More **complex architecture** than Mamba
- **Higher computational cost** than pure Mamba

Mamba & Samba – Key Features

➢ **Mamba vs. Samba – Comparison Table**

Feature	Mamba	Samba
Core Concept	State Space Model (SSM)	Mamba + Self-Attention
Self-Attention	❌ No	✅ Yes
Computational Complexity	O(N) (Highly Efficient)	O(N)~O(N^2) (Hybrid)
Long Sequence Handling	✅ Excellent	✅ Good
Short Sequence Handling	❌ Weaker	✅ Strong
Best Use Cases	Long text, audio, time-series	General NLP, variable-length sequences

Mamba & Samba – Key Features

☑ Choose Mamba if:

- You need to process **very long sequences efficiently**
- Memory and computation efficiency are key concerns

☑ Choose Samba if:

- You need **both long and short sequence performance**
- You require **self-attention for specific NLP tasks**

🚀 Future Outlook:

- Mamba and Samba continue to evolve, with potential applications in **AI research**, **time-series analysis**, and **NLP advancements**.