

# BESCHREIBUNG

## Bezeichnung:

Hardware-nahe Vorrichtung und Verfahren zur dynamischen Speicherpartitionierung und Modell-Transformation in adaptiven Systemen mittels kontextabhängiger Signalgatterung

## Technisches Gebiet

Die vorliegende Erfindung betrifft das Gebiet der adaptiven Signalverarbeitung und Hardware-nahen Steuerung von Speicherressourcen in Systemen der künstlichen Intelligenz, insbesondere für autoregressive Sprachmodelle (LLMs) und generative Diffusionsmodelle.

## Stand der Technik

Bisherige Systeme leiden beim Training unter „Katastrophalem Vergessen“. Bestehende Gatter-Mechanismen zur Steuerung des Signalflusses sind zudem rechenintensiv ( $O(n \cdot k)$ ) und unflexibel gegenüber strukturellen Modell-Erweiterungen.

## Darstellung der Erfindung

Die Erfindung löst diese Probleme durch eine asymmetrische Gatter-Logik. Ein **Kontext-Detektor (10)** analysiert zeitliche Zustandsfolgen des Eingangssignals. Dieser Detektor kann einstufig, mehrstufig oder in einer **hierarchischen Router-Architektur** organisiert sein, um komplexe Auswahlssignale für verschiedene Hierarchieebenen des Systems zu erzeugen.

Die Vorrichtung ermöglicht die Erweiterung adaptiver Systeme sowohl in der **Breite** (Erhöhung der Merkmalsdimensionen) als auch in der **Tiefe** (Hinzufügen zusätzlicher Verarbeitungsschichten/Layer), wobei die **Gatter-Logik (20)** eine kontrollierte Integration neuer Parameter ohne Beeinflussung der Basiskonfiguration gewährleistet.

Die Gatter-Steuerung nutzt eine rangreduzierte Broadcast-Addition ( $g^{row} \oplus g^{col}$ ), was die Komplexität auf  $O(n + k)$  reduziert. Eine integrierte **Rescue-Einheit (30)** reagiert bei Bedarf Pfade durch Signalinjektion. Die Signale werden in spezifische **Speicher-segmente (40)** geleitet, unterteilt in einen schreibgeschützten „Frozen Core“ und eine adaptive „Expansion Area“.

## Beschreibung der Zeichnungen

Fig. 1 zeigt die schematische Hardware-Architektur der Vorrichtung (10-40).

Fig. 2 verdeutlicht die rangreduzierte Gatter-Berechnung innerhalb der Logik (20).

# PATENTANSPRÜCHE

1. Signalverarbeitungsvorrichtung zur Steuerung von Speicherzugriffen, umfassend einen Kontext-Detektor (10) zur Erzeugung eines Auswahlsignals basierend auf zeitlichen Zustandsfolgen und eine asymmetrische Gatter-Logik (20) zur selektiven Schreibsperrung von Speichersegmenten (40).
2. Vorrichtung nach Anspruch 1, dadurch gekennzeichnet, dass der Kontext-Detektor (10) eine **hierarchische Struktur** aus mehreren logischen Router-Einheiten aufweist.
3. Vorrichtung nach Anspruch 1 oder 2, gekennzeichnet durch eine Broadcast-Recheneinheit zur rangreduzierten Gatter-Steuerung mittels Vektor-Addition.
4. Vorrichtung nach einem der Ansprüche 1 bis 3, gekennzeichnet durch eine Rescue-Einheit (30) zur aktiven Pfadreaktivierung mittels Signalinjektion.
5. Verfahren zur Modell-Transformation, wobei Speichersegmente (40) so angesteuert werden, dass ein Modell in der **Breite oder Tiefe** durch eine „Expansion Area“ erweitert wird, während die Basiskonfiguration als „Frozen Core“ schreibgeschützt bleibt.

# ZEICHNUNGEN

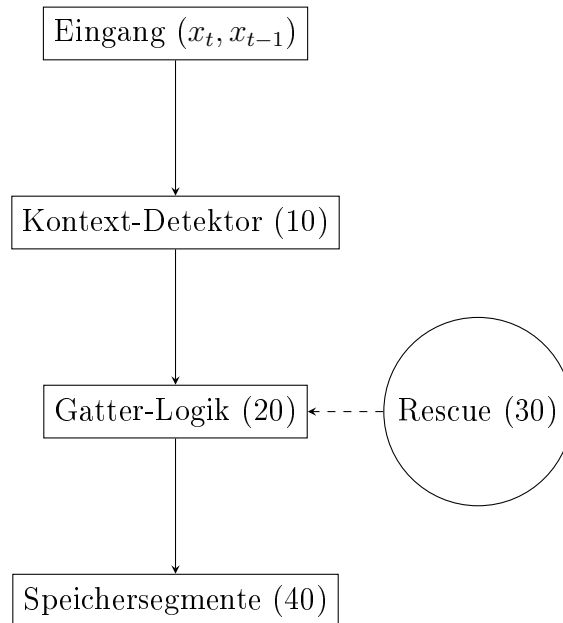


Abbildung 1: \*

Fig. 1: Systemarchitektur mit Kontext-Steuerung.

## ZUSAMMENFASSUNG

Die Erfindung betrifft eine hardware-nahe Vorrichtung zur effizienten Steuerung von Speicherressourcen. Ein Kontext-Detektor (10), der auch hierarchisch aufgebaut sein kann, erzeugt Auswahlsignale für eine asymmetrische Gatter-Logik (20). Diese nutzt rangreduzierte Berechnungen zur Ressourceneinsparung. Die Vorrichtung ermöglicht die Erweiterung von Modellen in Breite und Tiefe ohne Informationsverlust in bestehenden Speichersegmenten (40).