



Adaptive Resonance Theory

Jonas Jacobi, Felix J. Oppermann

C.v.O. Universität Oldenburg



Gliederung



1. Neuronale Netze
2. Stabilität - Plastizität
3. ART-1
4. ART-2
5. ARTMAP
6. Anwendung
7. Zusammenfassung



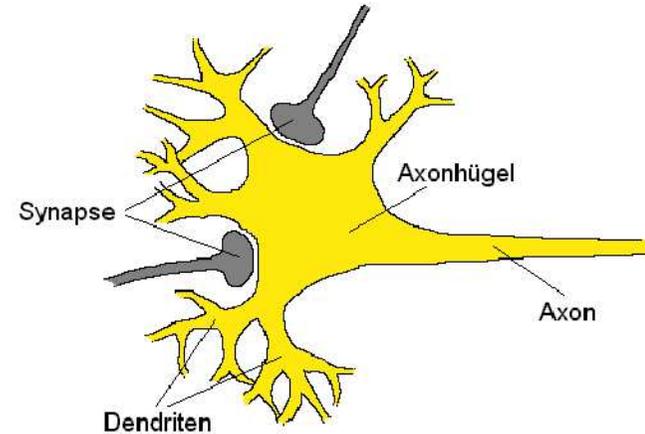
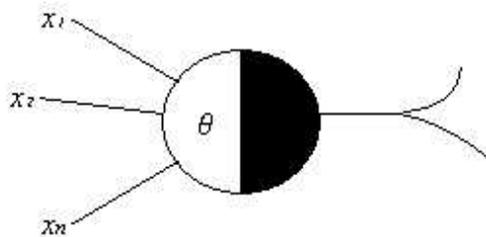
Neuronale Netze

- Künstliche neuronale Netze sind Modell für das Nervensystem
- Aufbau aus einzelnen Zellen
- Parallele Datenverarbeitung
- Lernen implizit durch Gewichtsveränderung
- Sehr unterschiedliche Strukturen

Neuronale Netze



- Summation der Eingänge
- Vergleich mit einem Schwellenwert
- Erzeugen eines Ausgabesignals



Stabilität \leftrightarrow Plastizität

- Lernen mit fester Lernrate
- Die Lernrate gibt an wie stark ein neues Muster die Gewichte beeinflusst
- Lernen erfolgt bei überwachtem Lernen nur in einer Trainingsphase
- Im biologischen System wird das Lernen nicht beendet
- Meist kein direktes Fehlersignal

Stabilität \leftrightarrow Plastizität



Stabilität

- Einmal Gelerntes soll gespeichert bleiben
- Das trainieren neuer Muster soll vermieden werden, da diese die Klassifikation stark verändern könnten

Plastizität

- Neue Muster sollen kein vollständiges Neutrainieren erfordern
- Alte Klassifikation sollten beim Hinzufügen neuer Muster erhalten bleiben



Stabilität \leftrightarrow Plastizität

- Beide Anforderungen konkurrieren
- Hohe Stabilität oder hohe Plastizität
- Plastizitäts-Stabilitäts-Dilemma

Stabilität \leftrightarrow Plastizität



Lösungsansätze:

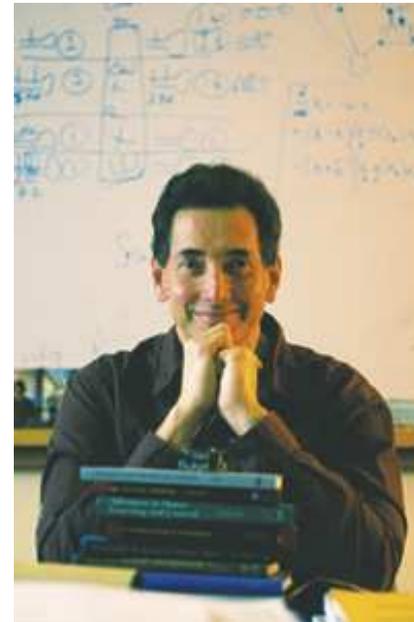
- Kontrolle ob das neue Muster bereits bekannt ist
- Für neue Muster eine neue Klasse öffnen
- Bei hoher Ähnlichkeit zu bekannten Mustern Einordnung in diese Klasse

Der Parameter der Kontrolle kann als Aufmerksamkeit betrachtet werden



Stephen Grossberg

- Stuyvesant High School, Manhattan, 1957
- Rockefeller University, Ph.D., 1964-1967
- Associate Professor of Applied Mathematics, M.I.T., 1969-1975.
- Professor of Mathematics, Psychology, and Biomedical Engineering, Boston University, 1975-heute



Adaptive Resonance Theory

- Von Grossberg 1976 entwickelte Theorie
- Versucht das Plastizitäts-Stabilitäts-Dilemma zu lösen
- Grundlage verschiedener neuronaler Modelle

Adaptive Resonance Theory



Resonance:

- Gegenseitige Verstärkung

Adaptive:

- Gewichts Anpassung nur bei Resonanz



ART-Netze

- Kombination überwachten und unüberwachten Lernens
- Neue Kategorien werden dynamisch angelegt
- Durch biologische Systeme motiviert
- Besitzen Kurz- und Langzeitgedächtnis

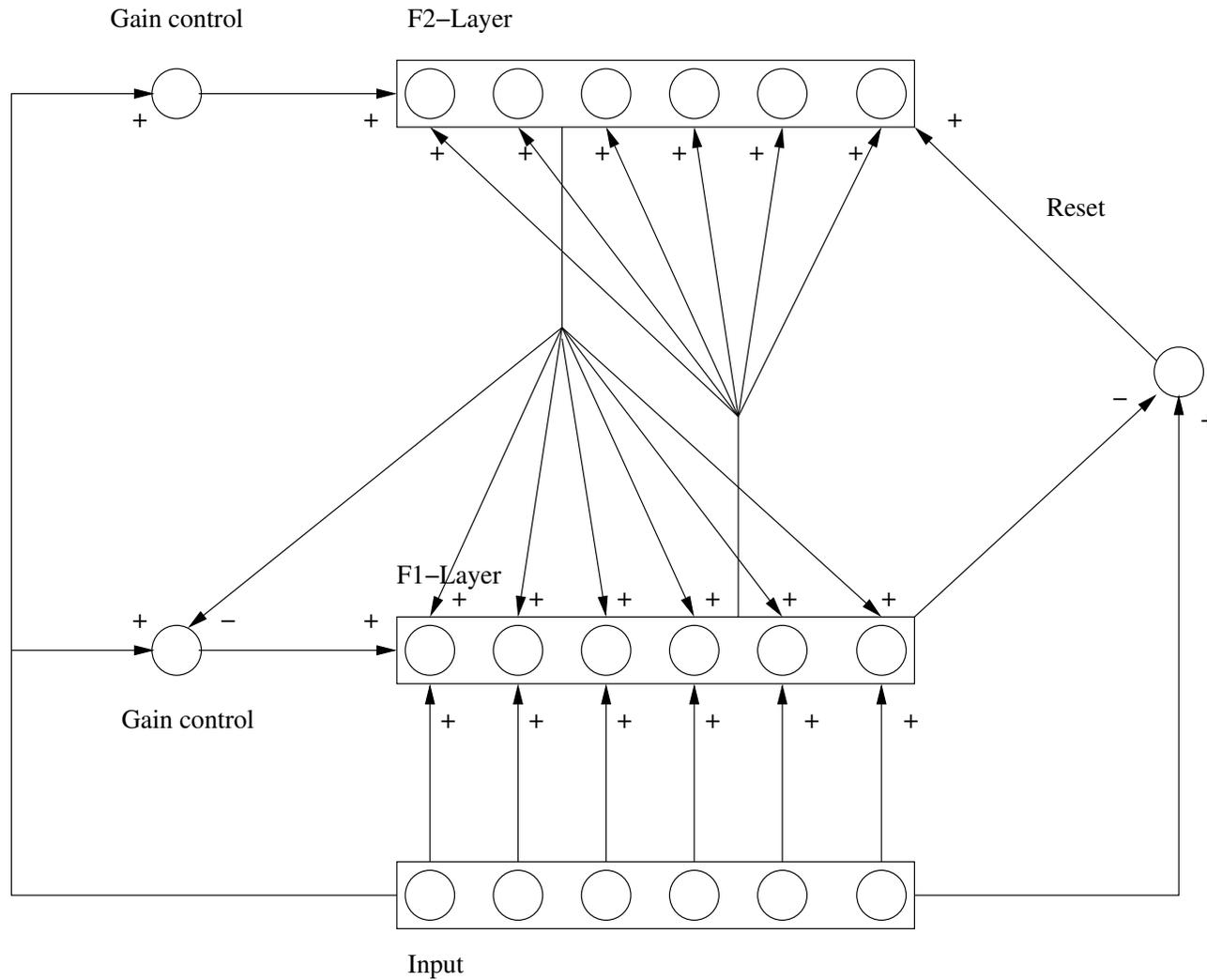
ART-1



- Von Stephen Grossberg 1976 vorgeschlagenes Modell
- Das ART-1 soll ein Modell kognitiver Prozesse sein
- Verarbeitet binäre Signale



ART-1 / Struktur



ART-1 / Struktur



F1-Schicht:

- Erhält die Eingabe
- Stellt ein Kurzzeitgedächtnis dar

F2-Schicht:

- Jedes Neuron repräsentiert eine Kategorie
- Langzeitgedächtnis in den Verbindungen zu F1-Schicht



ART-1 / Algorithmus

- Eingabe wird an F1-Schicht angelegt und zwischengespeichert
- Weiterleitung an F2-Schicht
- Hypothese über eine Klassifizierung in der F2-Schicht bilden
- Rückgabe auf Grundlage dieser Hypothese erzeugen
- Vergleich in der F1-Schicht. Falls die Hypothese nahe genug an der Eingabe ist: Anpassen der Gewichte

ART-1 / Algorithmus

- Sonst erzeugen eine neuen Hypothese unter Ausschluss der schon zuvor getesteten Klassen
- Wiederholung bis eine passende Klasse gefunden wurde
- Kann die Eingabe keiner Klasse zugeordnete werden, so wird eine neue Klasse erzeugt

ART-1 / Hilfsneuronen

- Die Neuronen des ART-1 müssen bei verschiedenen Schritten des Algorithmus verschieden reagieren
- Neuronen selbst besitzen keinen Mechanismus, der dies ermöglicht
- Gain control, G_1 und G_2 , und das Reset, R lösen dieses Problem
- Lösen ein rein technisches Problem

ART-1 / 2/3-Regel

- Es gibt aktivierende und hemmende Eingänge
- Die Neuronen der F1- und F2-Schicht feuern genau dann, wenn genau zwei ihrer drei Eingänge aktiv sind

ART-1 / Notation



n Komponentenanzahl des Eingabevektors.

m Maximale Klassenanzahl.

b_{ij} Gewichte von $Neuron_i$ in F_1 zu $Neuron_j$ in F_2 .

t_{ji} Gewichte von $Neuron_j$ in F_2 zu $Neuron_i$ in F_1 .

ρ Aufmerksamkeit.

s Eingabevektor.

x Aktivierungsvektor von F_1 .

$\|x\|$ Anzahl der Einsen des Vektors x .

L Lernrate



ART-1 / Lernalgorithmus

- 0 Initialisieren: $L > 1$, $0 < \rho \leq 1$, Anfangsgewichte $0 < b_{ij}(0) < \frac{L}{L-1+n}$ und $t_{ji} = 1$.
- 1 Durchlaufe Schritt **2-11** für jedes Eingabemuster.
 - 2 Setze die Aktivierung aller Neurone in F_2 auf 0. Setze die Aktivierung der Neurone von *Input* auf s .
 - 3 Berechne $\|s\|$.
 - 4 Gebe Eingangssignal von *Input* an F_1 weiter ($x_i = s_i$).
 - 5 Berechne $y_j = \sum_i b_{ij}x_i$ für alle Neurone in F_2 , die nicht deaktiviert sind.
 - 6 Solange reset gesetzt ist, führe Schritt **7-10** aus.

ART-1 / Lernalgorithmus



7 Setze

J = Index des Neurons in F_2 , das den höchsten y_j -Wert hat.

8 Berechne Aktivierung x von F_1 : $x_i = s_j t_{Ji}$.

9 Berechne $\|x\|$.

10 Test auf reset:

$$\begin{cases} \text{Deaktiviere Neuron } J \text{ und gehe zu Schritt 6,} & \text{wenn } \frac{\|x\|}{\|s\|} < \rho \\ \text{gehe zu Schritt 11,} & \text{sonst} \end{cases}$$

11 Erneure Gewichte des Neurons J : $b_{iJ} = \frac{Lx_i}{L-1+\|x\|}$,

$t_{Ji} = x_i$.



ART-2



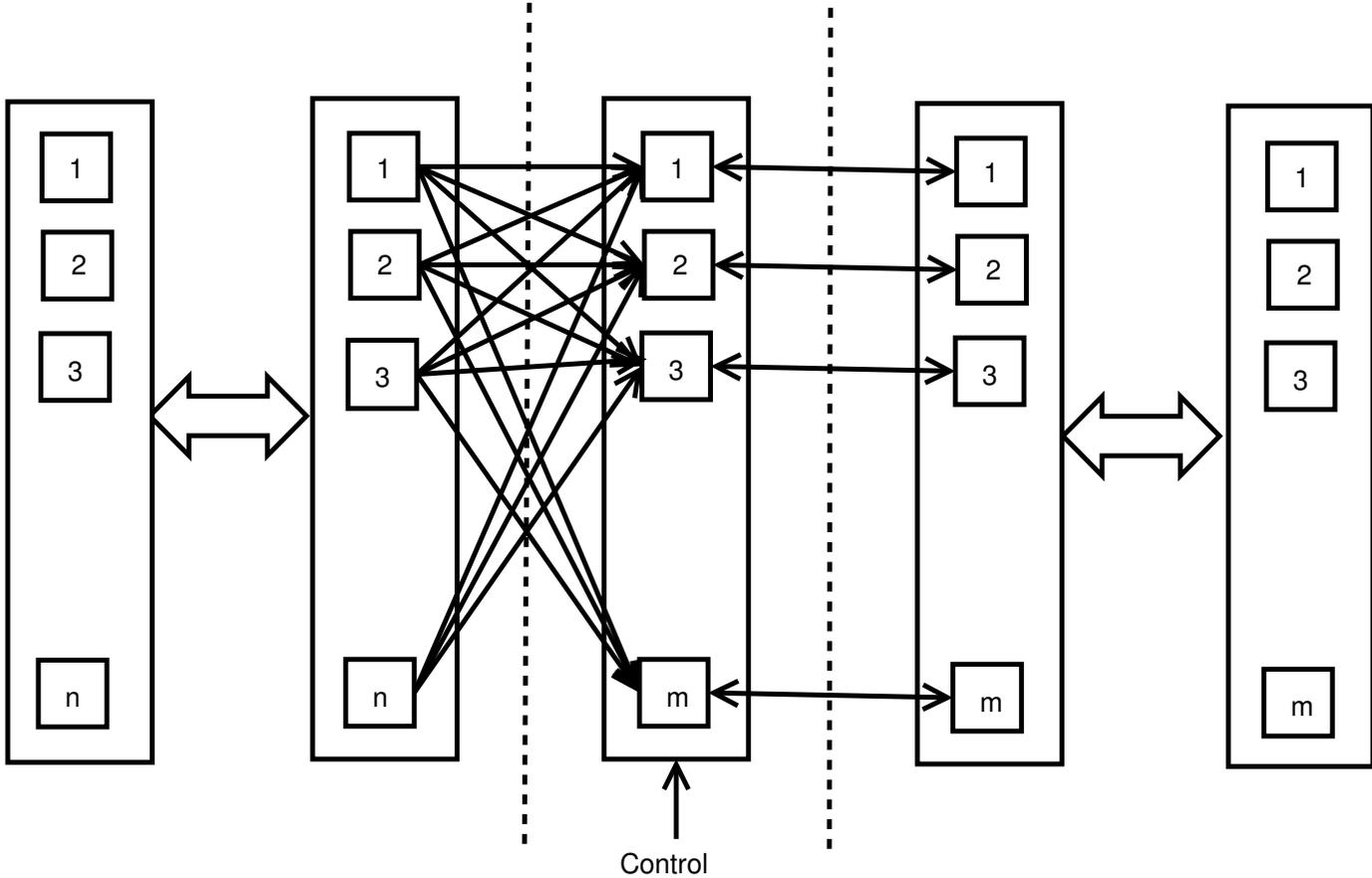
- Von Stephen Grosserg und Gail A. Carpenter 1987 entwickelte Modelle
- Verallgemeinerung der ART-Architektur
- Erlaubt reellwertige Eingaben und Gewichte
- Funktionsweise analog zu ART-1



ARTMAP

- Erweiterung um überwachtes Lernen zu ermöglichen.
- Besteht aus zwei ART-1 Netzen (ART_a , ART_b) und einem zwischengeschalteten Assoziationsnetz.
- Es wird eine Assoziation zwischen zwei Mustern a und b gebildet, die selbst Klasseneinteilungen von zwei Eingabemustern x_a und x_b sind.
- ART_b fungiert als „Lehrer“. Sollte die von ART_a vorhergesagte Klasse nicht mit der von ART_b übereinstimmen, wird das „Verhalten“ von ART_a angepasst.

ARTMAP



Anwendungen



- Modellierung biologischer neuronaler Vorgänge
- Musikerkennung
- Medizinische Diagnose
- E-Mail-Klassifizierung (Erkennen von Ähnlichkeiten bei Texten)



Zusammenfassung

- Die Adaptive Resonance Theory ist ein Beispiel für den Versuch kognitive Vorgänge durch neuronale Netze zu modellieren
- Zeigt Teilaspekte des Verhaltens des biologischen Vorbilds
- Bei weitem keine vollständige Modellierung
- Ein Forschungsansatz unter vielen
- Grundlage auch für technische Lösungen

Literatur



- [Bot98] Hans-Heinrich Bothe. *Neuro-Fuzzy-Methoden: Einführung in die Theorie und Anwendungen*. Springer, Berlin, 1998.
- [Bra95] Rüdiger Brause. *Neuronale Netze: Eine Einführung in die Neuroinformatik*. Teubner, Stuttgart, 1995.
- [Che00] Ke Chen. Lectureslides Nature Inspired Learning
(http://ailab.ch/teaching/classes/2004ss/nn/Lecture_on_ART.pdf),
2000.
- [Jen04] Peter Jensch. Skript zu Vorlesung Neuro-Fuzzy-Methoden. 2004.
- [Roj96] Raul Rojas. *Theorie der neuronalen Netze*. Springer, Berlin, Vierte Ausgabe, 1996.
- [RT98] Edmund T. Rolls und Alessandro Treves. *Neural Networks and Brain Function*. Oxford University Press, New York, 1998.

