شبکه‌های عصبی مصنوعی

درس ۱۹

نظریه‌ی تشدید وفقی

Adaptive Resonance Theory

کاظم فولادی قلعه

دانشکده مهندسی، پردیس فارابی

دانشگاه تهران

http://courses.fouladi.ir/nn

Adaptive Resonance Theory (ART)

نظریه‌ی تشدید وفقی

ADAPTIVE RESONANCE THEORY (ART)

یک مشکل کلیدی در شبکه‌های رقابتی: **پایداری یادگیری**:
هیچ تضمینی وجود ندارد که اِعمال همه‌ی ورودی‌ها منجر به همگرایی شبکه شود.

شبکه‌ی ART: یک نوع تغییریافته از شبکه‌های رقابتی
که به‌منظور غلبه بر مشکل پایداری یادگیری طراحی شده است.

نظریه‌ی تشدید وفقی

۱

مروری بر تشدید وفقی

نظریه‌ی تشدید وفقی

مشکل پایداری در شبکه‌های رقابتی

ADAPTIVE RESONANCE THEORY (ART)

مشکل اصلی شبکه‌ی گراسبرگ و سایر شبکه‌های رقابتی:
آنها همیشه قادر به ایجاد خوشه‌ها یا دسته‌های پایدار نیستند.

گراسبرگ نشان داد که:

- اگر تعداد الگوهای ورودی زیاد نباشد
  و یا الگوهای ورودی دارای تعداد خوشه‌ی زیادی (نسبت به تعداد نرون‌های لایه‌ی دوم) نباشند؛
  آن‌گاه یادگیری از پایداری مناسبی برخوردار خواهد بود.

- شبکه‌های رقابتی استاندارد در پاسخ به الگوهای ورودی **دلخواه** فاقد یادگیری پایدار هستند.

دلیل ناپایداری یادگیری شبکه‌های رقابتی: **قابلیت وفق‌پذیری (و انعطاف) آنها**
این قابلیت منجر به تأثیرپذیری بیشتر یادگیری از ورودی‌های تازه‌تر می‌شود.

گراسبرگ به این مشکل «معمای پایداری ـ انعطاف‌پذیری» می‌گوید.

نظریه‌ی تشدید وفقی

مشکل پایداری در شبکه‌های رقابتی: معمای پایداری ـ انعطاف‌پذیری

STABILITY/PLASTICITY DILEMMA

معمای پایداری ـ انعطاف‌پذیری:
سیستم چگونه می‌تواند علاوه بر تأثیرپذیری از نمونه‌های جدید
در تقابل با الگوهای نامربوط پایدار باقی بماند؟

سیستم‌های بیولوژیکی در این مورد بسیار خوب عمل می‌کنند:
برای مثال: انسان‌ها به‌آسانی قادر به تشخیص چهره‌ی مادرشان هستند،
حتی اگر برای مدتی طولانی او را ندیده باشند و در این مدت افراد بسیار زیادی را ملاقات کرده باشند.

در راستای حل این معما، گراسبرگ و کارپنتر یک نظریه‌ی جدید با عنوان
نظریه‌ی تشدید وفقی (ART)
را توسعه دادند.

شبکه‌های ART بر مبنای عملکرد شبکه‌های گراسبرگ می‌باشند؛
یک نوآوری کلیدی در ART استفاده از مفهوم «انتظار» (Expectation) است.

نظریه‌ی تشدید وفقی

مفهوم «انتظار»

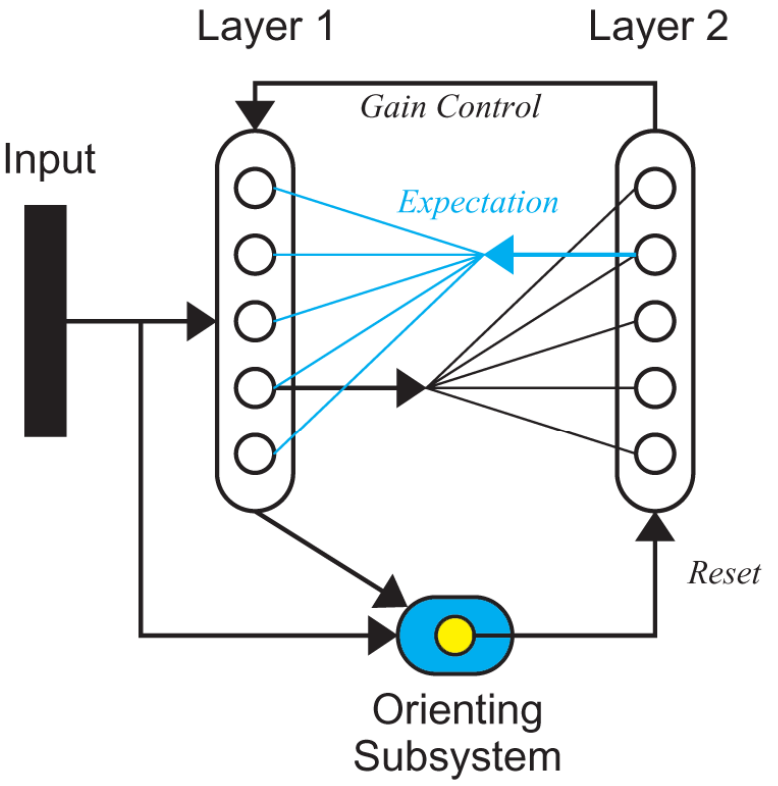
"EXPECTATION" CONCEPT

انتظار: با ارائه‌ی هر یک از الگوهای ورودی به شبکه،
آن ورودی با نزدیک‌ترین بردار پروتوتایپ به خود مقایسه می‌شود (انتظار).

اگر میزان نزدیکی بردار ورودی به بردار پروتوتایپ کافی نباشد، یک بردار پروتوتایپ دیگر انتخاب می‌شود.

⇐ الگوهای یادگرفته شده (الگوهای پروتوتایپ)، با توجه به یادگیری‌های جدیدتر دچار فرسایش نمی‌شوند.

Basic ART Architecture

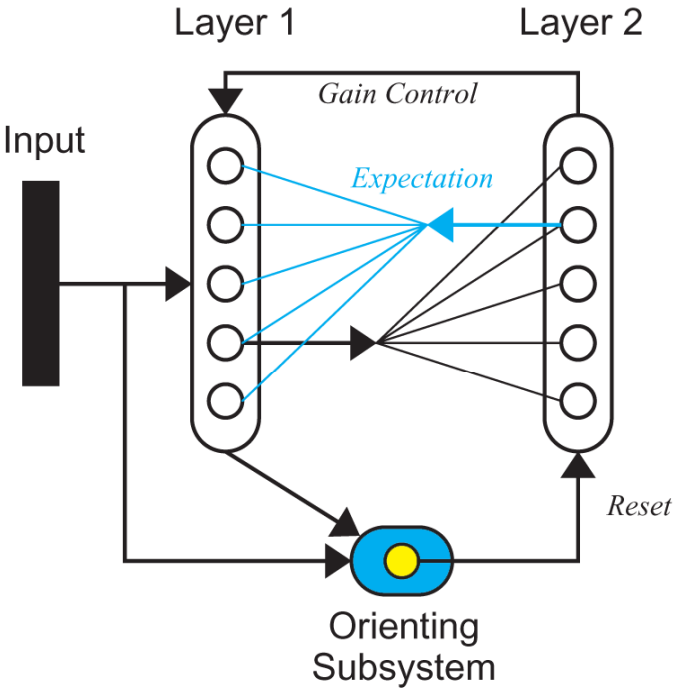


نظریه‌ی تشدید وفقی

معماری پایه‌ی ART

BASIC ART ARCHITECTURE

معماری ART1: این شبکه خاص ورودی‌های دودویی طراحی شده است.



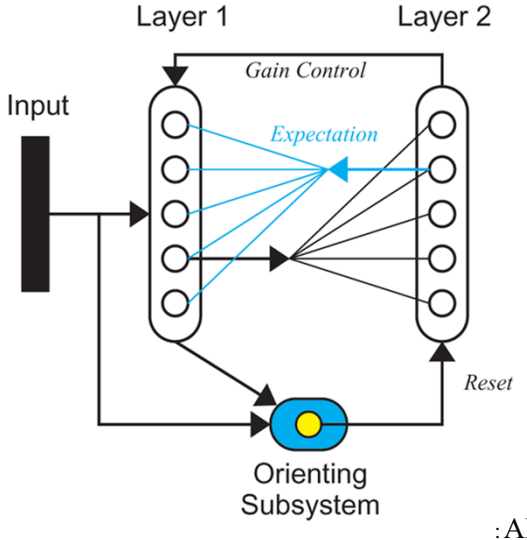
این شبکه، شکل تغییریافته‌ی شبکه‌ی گراسبرگ به‌منظور **پایدارسازی** فرآیند یادگیری است.

نظریه‌ی تشدید وفقی

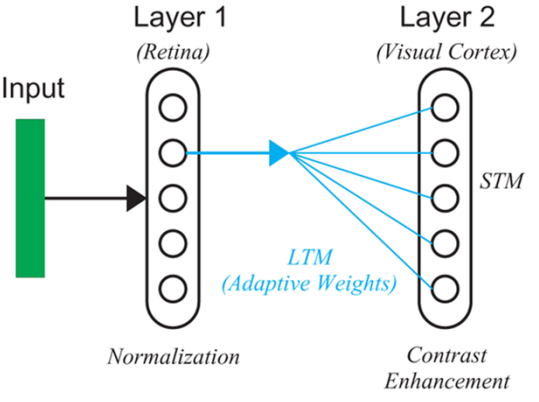
معماری پایه‌ی ART در مقایسه با گراسبرگ

BASIC ART ARCHITECTURE VS. GROSSBERG

ART



Grossberg



نوآوری‌ها در معماری ART1:

- انتظار لایه‌ی دوم به لایه‌ی اول
- زیرسیستم گرادهنده
- کنترل بهره

نظریه‌ی تشدید وفقی

اتصالات لایه‌ها: لایه‌ی ۱ به ۲

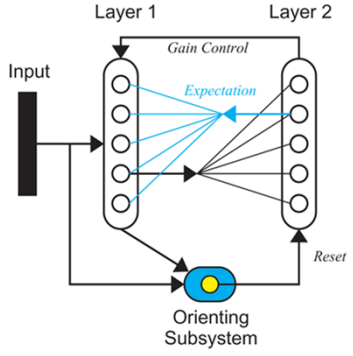
L1-L2 Connections

اتصالات لایه‌ی ۱ به ۲ در شبکه‌های گراسبرگ (و ART) از نوع Instar است.
این اتصالات عملیات خوشه‌بندی را انجام می‌دهند.

- وقتی یک الگوی ورودی به شبکه اعمال می‌شود، پس از نرمال‌سازی در ماتریس وزن L1-L2 ضرب می‌شود.
- سپس در لایه‌ی ۲ با انجام عملیات رقابت، سطری از ماتریس وزن که در نزدیک‌ترین وضعیت نسبت به بردار ورودی قرار دارد، تشخیص داده می‌شود.
- سپس این سطر به سمت بردار ورودی جابه‌جا می‌شود.

⇓

پس از تکمیل یادگیری، هر یک از سطرهای ماتریس وزن L1-L2 حاوی یکی از بردارهای پروتوتایپ خواهد بود که بازنمایی کننده‌ی یک خوشه از بردارهای ورودی است.



نظریه‌ی تشدید وفقی

اتصالات لایه‌ها: لایه‌ی ۲ به ۱

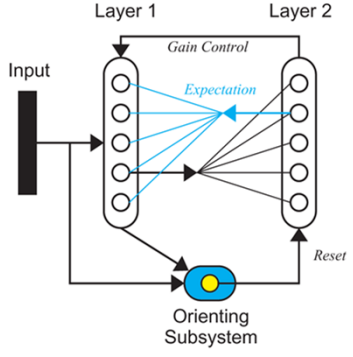
L2-L1 Connections

در شبکه‌های ART روند یادگیری در مجموعه‌ای از اتصالات فیدبک از لایه‌ی ۲ به لایه‌ی ۱ نیز رخ می‌دهد.

اتصالات لایه‌ی ۲ به ۱ در شبکه‌های ART از نوع Outstar است.
این اتصالات عملیات به‌یادآوری الگوها را انجام می‌دهند.

- هرگاه یک گره در لایه‌ی ۲ فعال شود، فیدبک، الگوی ورودی در لایه‌ی ۱ را بازتولید می‌کند **(الگوی انتظار)**.
- سپس لایه‌ی ۱ یک مقایسه بین الگوی انتظار و الگوی ورودی انجام می‌دهد.
- اگر الگوی انتظار و الگوی ورودی، به‌اندازه‌ی کافی به هم نزدیک نباشند، زیرسیستم گرادهنده یک reset در لایه‌ی ۲ ایجاد می‌کند: این reset نرون برنده‌ی فعلی را غیرفعال کرده و مقدار انتظار جاری را حذف می‌کند.
- سپس یک رقابت جدید در لایه‌ی ۲ بدون حضور نرون برنده‌ی قبلی انجام می‌شود.
- نرون برنده‌ی جدید در لایه‌ی ۲ مقدار انتظار جدید را به لایه‌ی ۱ منعکس می‌کند (از طریق اتصالات L2-L1).

این فرآیند آن‌قدر تکرار می‌شود تا انتظار L2-L1 به‌اندازه‌ی کافی به الگوی ورودی نزدیک شود.



ART Subsystems

Layer 1

- Normalization
- Comparison of input pattern and expectation

L1-L2 Connections (Instars)

- Perform clustering operation.
- Each row of $W^{1:2}$ is a prototype pattern.

Layer 2

- Competition, contrast enhancement

L2-L1 Connections (Outstars)

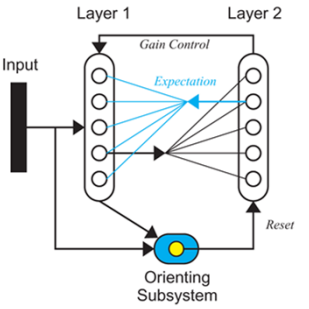
- Expectation
- Perform pattern recall.
- Each column of $W^{2:1}$ is a prototype pattern

Orienting Subsystem

- Causes a reset when expectation does not match input
- Disables current winning neuron

نظریه‌ی تشدید وفقی

ART Subsystems
زیرسیستم‌های ART



لایه‌ی ۱
- نرمال‌سازی
- مقایسه‌ی الگوی ورودی و انتظار

اتصالات لایه‌ی ۲ به ۱ (Instar)
- انجام عملیات خوشه‌بندی
- هر سطر $\mathbf{W}^{1:2}$ یک الگوی پروتوتایپ است.

لایه‌ی ۲
- رقابت، ارتقای کنتراست

اتصالات لایه‌ی ۱ به ۲ (Outstar)
- انتظار
- انجام عملیات به‌یادآوری الگوها
- هر ستون $\mathbf{W}^{2:1}$ یک الگوی پروتوتایپ است.

زیرسیستم گرادهنده
- ایجاد یک reset هنگام عدم تطبیق انتظار با ورودی
- غیرفعال‌سازی نرون برنده‌ی فعلی

Layer 1
- Normalization
- Comparison of input pattern and expectation

L1-L2 Connections (Instars)
- Perform clustering operation.
- Each row of $W^{1:2}$ is a prototype pattern.

Layer 2
- Competition, contrast enhancement

L2-L1 Connections (Outstars)
- Expectation
- Perform pattern recall.
- Each column of $W^{2:1}$ is a prototype pattern

Orienting Subsystem
- Causes a reset when expectation does not match input
- Disables current winning neuron

شبکه‌های عصبی مصنوعی

نظریه‌ی تشدید وفقی

۲

لایه‌ی ۱

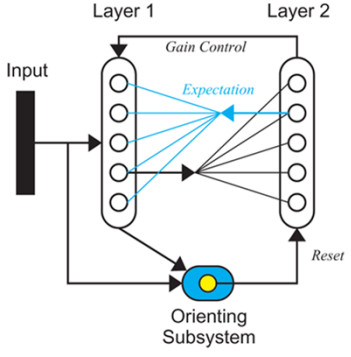
نظریه‌ی تشدید وفقی

لایه‌ی ۱

LAYER 1

هدف اصلی لایه‌ی ۱، مقایسه‌ی الگوی ورودی با الگوی انتظار از لایه‌ی ۲ می‌باشد.
(هر دوی این الگوها در ART1 دودویی هستند.)

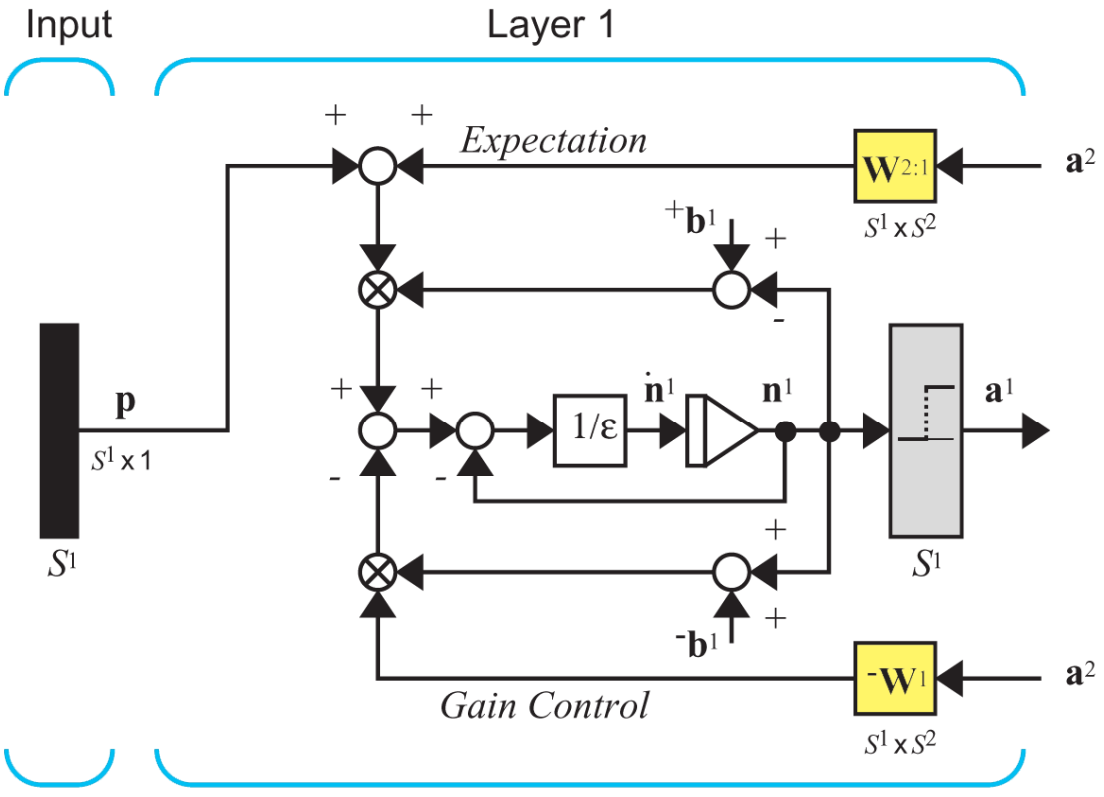
- اگر این دو الگو، به‌اندازه‌ی کافی به هم نزدیک نباشند،
  آنگاه: زیرسیستم گرادهنده یک reset در لایه‌ی ۲ ایجاد می‌کند.
- اگر این دو الگو، به‌اندازه‌ی کافی به هم نزدیک باشند،
  آنگاه: لایه‌ی ۱ الگوی انتظار و ورودی را با هم ترکیب کرده و یک الگوی پروتوتایپ جدید تولید می‌کند.



لایه‌ی ۱ شبکه‌ی ART1 بسیار شبیه به لایه‌ی ۱ شبکه‌ی گراسبرگ است.

تفاوت: در ورودی‌های تهییج‌گر و بازدارنده‌ی مدل شانتینگ است:
در شبکه‌ی ART1 هیچگونه نرمال‌سازی در لایه‌ی ۱ انجام نمی‌شود؛
⇓
نیازی به اتصالات «مرکز–روشن/پیرامون–خاموش»
بین بردار ورودی و لایه‌ی ۱ وجود ندارد.

Layer 1

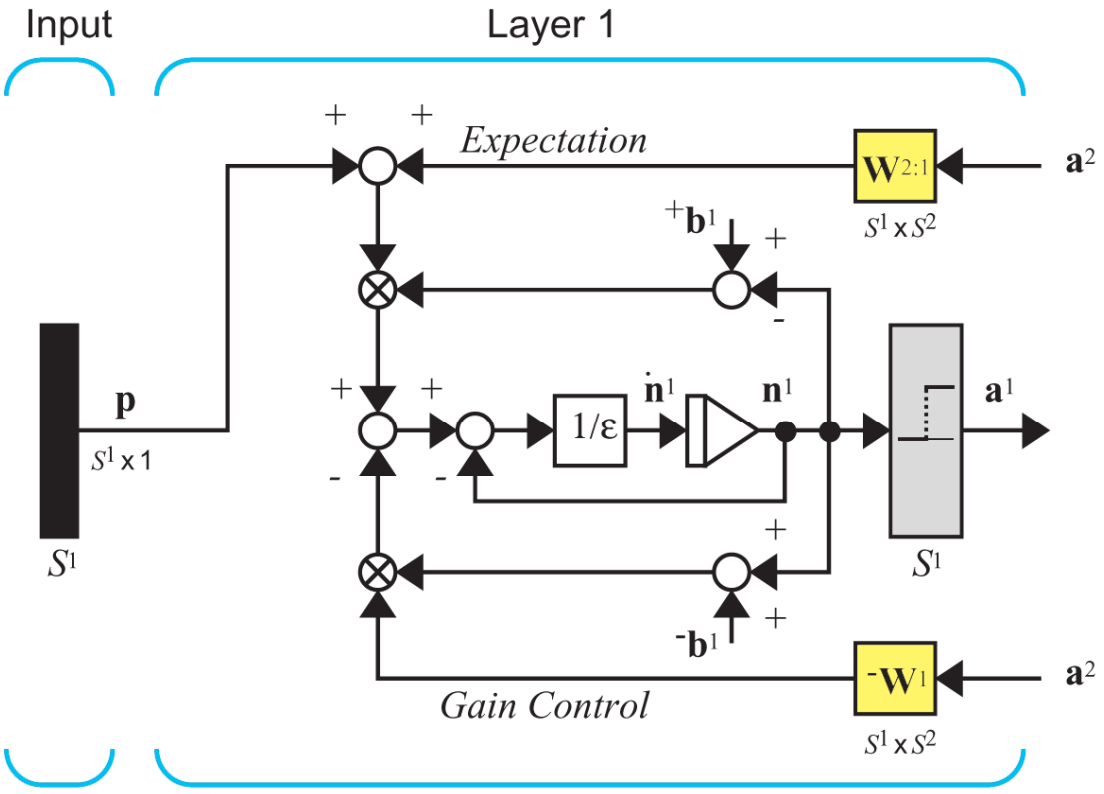


$$\varepsilon \, d\mathbf{n}^1/dt = -\mathbf{n}^1 + (^{+}\mathbf{b}^1 - \mathbf{n}^1)\{\mathbf{p} + \mathbf{W}^{2:1}\mathbf{a}^2\} - (\mathbf{n}^1 + {}^{-}\mathbf{b}^1)[{}^{-}\mathbf{W}^1]\mathbf{a}^2$$

نظریه‌ی تشدید وفقی

لایه‌ی ۱

LAYER 1



$$\varepsilon \, d\mathbf{n}^1/dt = -\mathbf{n}^1 + (^{+}\mathbf{b}^1 - \mathbf{n}^1)\{\mathbf{p} + \mathbf{W}^{2:1}\mathbf{a}^2\} - (\mathbf{n}^1 + {}^{-}\mathbf{b}^1)[{}^{-}\mathbf{W}^1]\mathbf{a}^2$$

Prepared by Kazim Fouladi | Fall 2017 | 2nd Edition
کاظم فولادی

Layer 1 Operation

Shunting Model

$$\varepsilon \frac{d\mathbf{n}^1(t)}{dt} = -\mathbf{n}^1(t) + ({}^{+}\mathbf{b}^1 - \mathbf{n}^1(t))\{\underbrace{\mathbf{p} + \mathbf{W}^{2:1}\mathbf{a}^2(t)}_{\text{Excitatory Input}}\} - (\mathbf{n}^1(t) + {}^{-}\mathbf{b}^1)[\underbrace{{}^{-}\mathbf{W}^1]\mathbf{a}^2(t)}_{\text{Inhibitory Input}}$$

Excitatory Input
(Comparison with Expectation)

Inhibitory Input
(Gain Control)

$$\mathbf{a}^1 = \mathbf{hardlim}^{+}(\mathbf{n}^1)$$

$$hardlim^{+}(n) = \begin{cases} 1, & n > 0 \\ 0, & n \le 0 \end{cases}$$

نظریه‌ی تشدید وفقی

عملیات لایه‌ی ۱

LAYER 1 OPERATION

Shunting Model

$$\varepsilon \frac{d\mathbf{n}^1(t)}{dt} = -\mathbf{n}^1(t) + ({}^{+}\mathbf{b}^1 - \mathbf{n}^1(t))\{\underbrace{\mathbf{p} + \mathbf{W}^{2:1}\mathbf{a}^2(t)}_{\text{Excitatory Input}}\} - (\mathbf{n}^1(t) + {}^{-}\mathbf{b}^1)[\underbrace{{}^{-}\mathbf{W}^1]\mathbf{a}^2(t)}_{\text{Inhibitory Input}}$$

Excitatory Input
(Comparison with Expectation)

Inhibitory Input
(Gain Control)

ورودی تهییج‌گر لایه‌ی ۱ شبکه‌ی ART1 شامل ترکیبی از الگوی ورودی و انتظار L1-L2 است.
ورودی بازدارنده‌ی لایه‌ی ۱ شبکه‌ی ART1 از سیگنال کنترل بهره‌ی لایه‌ی ۲ تشکیل می‌شود.

$$\mathbf{a}^1 = \mathbf{hardlim}^{+}(\mathbf{n}^1)$$

خروجی لایه‌ی ۱

$$hardlim^{+}(n) = \begin{cases} 1, & n > 0 \\ 0, & n \le 0 \end{cases}$$

Excitatory Input to Layer 1

$$\mathbf{p} + \mathbf{W}^{2:1}\mathbf{a}^{2}(t)$$

Suppose that neuron j in Layer 2 has won the competition:

$$\mathbf{W}^{2:1}\mathbf{a}^{2} = \begin{bmatrix} \mathbf{w}_1^{2:1} & \mathbf{w}_2^{2:1} & \cdots & \mathbf{w}_j^{2:1} & \cdots & \mathbf{w}_{S^2}^{2:1} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \end{bmatrix} = \mathbf{w}_j^{2:1} \qquad (j\text{th column of } \mathbf{W}^{2:1})$$

Therefore the excitatory input is the sum of the input pattern and the L2-L1 expectation:

$$\mathbf{p} + \mathbf{W}^{2:1}\mathbf{a}^{2} = \mathbf{p} + \mathbf{w}_j^{2:1}$$

نظریه‌ی تشدید وفقی

ورودی تهییج‌گر به لایه‌ی ۱

EXCITATORY INPUT TO LAYER 1

ورودی تهییج‌گر لایه‌ی ۱ برابر با مجموع بردار ورودی و انتظار L2-L1 است.

$$\mathbf{p} + \mathbf{W}^{2:1}\mathbf{a}^{2}(t)$$

به‌عنوان مثال: فرض کنید که نرون j–ام در لایه‌ی ۲ رقابت را برده باشد ⇐ خروجی آن 1 و خروجی سایر نرون‌ها 0 است.

Suppose that neuron j in Layer 2 has won the competition:

$$\mathbf{W}^{2:1}\mathbf{a}^{2} = \begin{bmatrix} \mathbf{w}_1^{2:1} & \mathbf{w}_2^{2:1} & \cdots & \mathbf{w}_j^{2:1} & \cdots & \mathbf{w}_{S^2}^{2:1} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \end{bmatrix} = \mathbf{w}_j^{2:1} \qquad (j\text{th column of } \mathbf{W}^{2:1})$$

پس ورودی تهییج‌گر لایه ۱ برابر با مجموع الگوی ورودی و انتظار L2-L1 خواهد بود.

Therefore the excitatory input is the sum of the input pattern and the L2-L1 expectation:

$$\mathbf{p} + \mathbf{W}^{2:1}\mathbf{a}^{2} = \mathbf{p} + \mathbf{w}_j^{2:1}$$

هر ستون از ماتریس L2-L1 بازنمایی‌کننده‌ی یک انتظار متفاوت (الگوی پروتوتایپ) است.
لایه‌ی ۱، الگوی ورودی را با انتظار ترکیب می‌کند (با استفاده از عملیات AND منطقی).

Inhibitory Input to Layer 1

Gain Control

$$[-\mathbf{W}^1]\mathbf{a}^2(t)$$

$$-\mathbf{W}^1 = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}$$

The gain control will be one when Layer 2 is active (one neuron has won the competition), and zero when Layer 2 is inactive (all neurons having zero output).

نظریه‌ی تشدید وفقی

ورودی بازدارنده به لایه‌ی ۱

INHIBITORY INPUT TO LAYER 1

Gain Control

ورودی بازدارنده لایه‌ی ۱ برابر با جمله‌ی کنترل بهره است:

$$[\bar{\mathbf{W}}^1]\mathbf{a}^2(t)$$

$$\bar{\mathbf{W}}^1 = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}$$

کاربرد کنترل بهره؟ پس از آشنایی با تحلیل حالت ماندگار لایه‌ی ۱

The gain control will be one when Layer 2 is active (one neuron has won the competition), and zero when Layer 2 is inactive (all neurons having zero output).

بنابراین، ورودی بازدارنده به هر یک از نرون‌های لایه‌ی ۱ برابر با مجموع همه‌ی خروجی‌های لایه‌ی ۲ خواهد بود.
چون از رقابت «برنده همه را می‌خورد» در لایه‌ی ۲ استفاده می‌کنیم، فقط یک عضو از $\mathbf{a}^2$ غیر صفر خواهد بود.
پس ورودی کنترل بهره در لایه‌ی ۱ در صورت فعال بودن لایه‌ی ۲ برابر با 1 و در صورت غیر فعال بودن آن 0 خواهد بود.

Steady State Analysis: Case I

$$\varepsilon \frac{dn_i^1}{dt} = -n_i^1 + ({}^{+}b^1 - n_i^1)\left\{p_i + \sum_{j=1}^{S^2} w_{i,j}^{2:1} a_j^2\right\} - (n_i^1 + {}^{-}b^1) \sum_{j=1}^{S^2} a_j^2$$

Case I: Layer 2 inactive (each $a^2_j = 0$)

$$\varepsilon \frac{dn_i^1}{dt} = -n_i^1 + ({}^{+}b^1 - n_i^1)\{p_i\}$$

In steady state:

$$0 = -n_i^1 + ({}^{+}b^1 - n_i^1)p_i = -(1 + p_i)n_i^1 + {}^{+}b^1 p_i \quad \Rightarrow \quad n_i^1 = \frac{{}^{+}b^1 p_i}{1 + p_i}$$

Therefore, if Layer 2 is inactive:

$$\mathbf{a}^1 = \mathbf{p}$$

نظریه‌ی تشدید وفقی

تحلیل حالت ماندگار: مورد «الف»

STEADY STATE ANALYSIS: CASE I

پاسخ نرون i–ام از لایه‌ی ۱ به صورت زیر می‌باشد:

$$\varepsilon \frac{dn_i^1}{dt} = -n_i^1 + ({}^+b^1 - n_i^1)\left\{p_i + \sum_{j=1}^{S^2} w_{i,j}^{2:1} a_j^2\right\} - (n_i^1 + {}^-b^1)\sum_{j=1}^{S^2} a_j^2$$

که $\varepsilon << 1$ است ⇐ حافظه‌ی کوتاه‌مدت (خروجی نرون) بسیار سریع‌تر از حافظه‌ی بلندمدت (ماتریس وزن) تغییر می‌کند.

مورد «الف»: Case I: Layer 2 inactive (each $a^2_j = 0$) **لایه‌ی ۲ غیرفعال است**

$$\varepsilon \frac{dn_i^1}{dt} = -n_i^1 + ({}^+b^1 - n_i^1)\{p_i\}$$

In steady state: مشتق صفر

$$0 = -n_i^1 + ({}^+b^1 - n_i^1)p_i = -(1+p_i)n_i^1 + {}^+b^1 p_i \quad \Rightarrow \quad n_i^1 = \frac{{}^+b^1 p_i}{1+p_i}$$

پس اگر $p_i = 0$ باشد آن‌گاه $n_i^1 = 0$ است و اگر $p_i = 1$ باشد آن‌گاه $n_i^1 = {}^+b^1/2 > 0$ خواهد بود.

Therefore, if Layer 2 is inactive:

$$\mathbf{a}^1 = \mathbf{p}$$

چون تابع انتقال لایه‌ی ۱، hardlim^+ انتخاب شده است، پس $\mathbf{a}^1 = \mathbf{p}$.

⇐ هرگاه لایه‌ی۲ غیرفعال باشد، خروجی لایه‌ی ۱ برابر با الگوی ورودی خواهد بود.

Steady State Analysis: Case II

Case II: Layer 2 active (one $a^2_j = 1$)

$$\varepsilon \frac{dn_i^1}{dt} = -n_i^1 + ({}^{+}b^1 - n_i^1)\{p_i + w_{i,j}^{2:1}\} - (n_i^1 + {}^{-}b^1)$$

In steady state:

$$\begin{aligned} 0 &= -n_i^1 + ({}^{+}b^1 - n_i^1)\{p_i + w_{i,j}^{2:1}\} - (n_i^1 + {}^{-}b^1) \\ &= -(1 + p_i + w_{i,j}^{2:1} + 1)n_i^1 + ({}^{+}b^1(p_i + w_{i,j}^{2:1}) - {}^{-}b^1) \end{aligned} \quad \Rightarrow \quad n_i^1 = \frac{{}^{+}b^1(p_i + w_{i,j}^{2:1}) - {}^{-}b^1}{2 + p_i + w_{i,j}^{2:1}}$$

We want Layer 1 to combine the input vector with the expectation from Layer 2, using a logical AND operation:

$n^1{}_i$<0, if either $w^{2:1}{}_{i,j}$ or p_i is equal to zero.

$n^1{}_i$>0, if both $w^{2:1}{}_{i,j}$ or p_i are equal to one.

$$\left.\begin{array}{l} {}^{+}b^1(2) - {}^{-}b^1 > 0 \\ {}^{+}b^1 - {}^{-}b^1 < 0 \end{array}\right\} \quad {}^{+}b^1(2) > {}^{-}b^1 > {}^{+}b^1$$

Therefore, if Layer 2 is active, and the biases satisfy these conditions:

$$\mathbf{a}^1 = \mathbf{p} \cap \mathbf{w}_j^{2:1}$$

نظریه‌ی تشدید وفقی

تحلیل حالت ماندگار: مورد «ب»

STEADY STATE ANALYSIS: CASE I

فرض می‌کنیم نرون j–ام برنده‌ی لایه‌ی ۲ باشد:

مورد «ب»: Case II: Layer 2 active (one $a^2_j = 1$) **لایه‌ی ۲ فعال است**

$$\varepsilon \frac{dn_i^1}{dt} = -n_i^1 + ({}^{+}b^1 - n_i^1)\{p_i + w_{i,j}^{2:1}\} - (n_i^1 + {}^{-}b^1)$$

(خروجی یک نرون 1 و خروجی بقیه‌ی نرون‌ها 0 است.)

In steady state: مشتق صفر

$$0 = -n_i^1 + ({}^{+}b^1 - n_i^1)\{p_i + w_{i,j}^{2:1}\} - (n_i^1 + {}^{-}b^1)$$
$$= -(1 + p_i + w_{i,j}^{2:1} + 1)n_i^1 + ({}^{+}b^1(p_i + w_{i,j}^{2:1}) - {}^{-}b^1)$$
$$\Rightarrow \quad n_i^1 = \frac{{}^{+}b^1(p_i + w_{i,j}^{2:1}) - {}^{-}b^1}{2 + p_i + w_{i,j}^{2:1}}$$

لایه‌ی ۱ باید بردار ورودی را با انتظار لایه‌ی ۲ (بیان شده با $w_j^{2:1}$) ترکیب نماید [الگوهای دودویی ← AND]

We want Layer 1 to combine the input vector with the expectation from Layer 2, using a logical AND operation:

$n^1{}_i$<0, if either $w^{2:1}{}_{i,j}$ or p_i is equal to zero.

$n^1{}_i$>0, if both $w^{2:1}{}_{i,j}$ or p_i are equal to one.

$$\left.\begin{array}{l} {}^{+}b^1(2) - {}^{-}b^1 > 0 \\ {}^{+}b^1 - {}^{-}b^1 < 0 \end{array}\right\} \quad {}^{+}b^1(2) > {}^{-}b^1 > {}^{+}b^1$$

برای مثال، می‌توان قرار داد: ${}^{+}b^1 = 1$ و ${}^{-}b^1 = 1.5$

Therefore, if Layer 2 is active, and the biases satisfy these conditions:

$$\mathbf{a}^1 = \mathbf{p} \cap \mathbf{w}_j^{2:1}$$

در صورت ارضای شرایط فوق برای بایاس‌ها و فعال بودن لایه‌ی ۲، خروجی لایه‌ی ۱ به این صورت می‌شود.

برای پیاده‌سازی عملیات AND به سیگنال «کنترل بهره» نیاز داریم.

نظریه‌ی تشدید وفقی

استفاده از سیگنال «کنترل بهره» برای پیاده‌سازی عملیات AND

$$n_i^1 = \frac{{}^{+}b^1(p_i + w_{i,j}^{2:1}) - {}^{-}b^1}{2 + p_i + w_{i,j}^{2:1}}$$

صورت کسر خروجی n_i^1 را در نظر می‌گیریم:

جمله‌ی ${}^{-}b^1$ در جمله‌ی مربوط به کنترل بهره ضرب می‌شود.
در این مورد کنترل بهره برابر با 1 است.

در صورت عدم وجود این جمله و در صورت بزرگ‌تر از صفر بودن p_i یا $w_{i,j}^{2:1}$
حاصل معادله‌ی فوق و درنتیجه n_i^1 بزرگ‌تر از صفر خواهد بود.

⇓

به‌جای یک عملیات AND، یک عملیات OR خواهیم داشت.
(مطابق بحثی که بخش زیرسیستم گرادهنده صورت می‌گیرد، انجام AND در لایه‌ی ۱ ضروری است).

زمانی که لایه‌ی ۲ غیرفعال است، جمله‌ی مربوط به کنترل بهره صفر است.
(لازم است زیرا در این حالت انتظار داریم که لایه‌ی ۱ تنها به الگوی ورودی عکس‌العمل نشان دهد،
زیرا هیچ انتظاری توسط لایه‌ی ۲ فعال نمی‌شود.)

Layer 1 Summary

If Layer 2 is inactive (each $a^2_j = 0$)

$$\mathbf{a}^1 = \mathbf{p}$$

If Layer 2 is active (one $a^2_j = 1$)

$$\mathbf{a}^1 = \mathbf{p} \cap \mathbf{w}_j^{2:1}$$

نظریه‌ی تشدید وفقی

خلاصه‌ی لایه‌ی ۱

LAYER 1 SUMMARY

If Layer 2 is inactive (each $a^2_j = 0$)

$$\mathbf{a}^1 = \mathbf{p}$$

If Layer 2 is active (one $a^2_j = 1$)

$$\mathbf{a}^1 = \mathbf{p} \cap \mathbf{w}^{2:1}_j$$

Layer 1 Example

$$\varepsilon = 1,\ {}^{+}b^1 = 1 \text{ and } {}^{-}b^1 = 1.5 \qquad \mathbf{W}^{2:1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \qquad \mathbf{p} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Assume that Layer 2 is active, and neuron 2 won the competition.

$$(0.1)\frac{dn_1^1}{dt} = -n_1^1 + (1 - n_1^1)\{p_1 + w_{1,2}^{2:1}\} - (n_1^1 + 1.5)$$
$$= -n_1^1 + (1 - n_1^1)\{0 + 1\} - (n_1^1 + 1.5) = -3n_1^1 - 0.5$$
$$\Rightarrow \frac{dn_1^1}{dt} = -30n_1^1 - 5$$

$$(0.1)\frac{dn_2^1}{dt} = -n_2^1 + (1 - n_2^1)\{p_2 + w_{2,2}^{2:1}\} - (n_2^1 + 1.5)$$
$$= -n_2^1 + (1 - n_2^1)\{1 + 1\} - (n_2^1 + 1.5) = -4n_2^1 + 0.5$$
$$\Rightarrow \frac{dn_2^1}{dt} = -40n_2^1 + 5$$

نظریه‌ی تشدید وفقی

لایه‌ی ۱ : مثال

LAYER 1 EXAMPLE

فرض می‌کنیم دو نرون در لایه‌ی ۲ و دو عنصر در ورودی داشته باشیم:

$$\varepsilon = 1,\ {}^{+}b^1 = 1 \ \text{ and } \ {}^{-}b^1 = 1.5 \qquad \mathbf{W}^{2:1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \qquad \mathbf{p} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

ماتریس وزن — ورودی

Assume that Layer 2 is active, and neuron 2 won the competition.

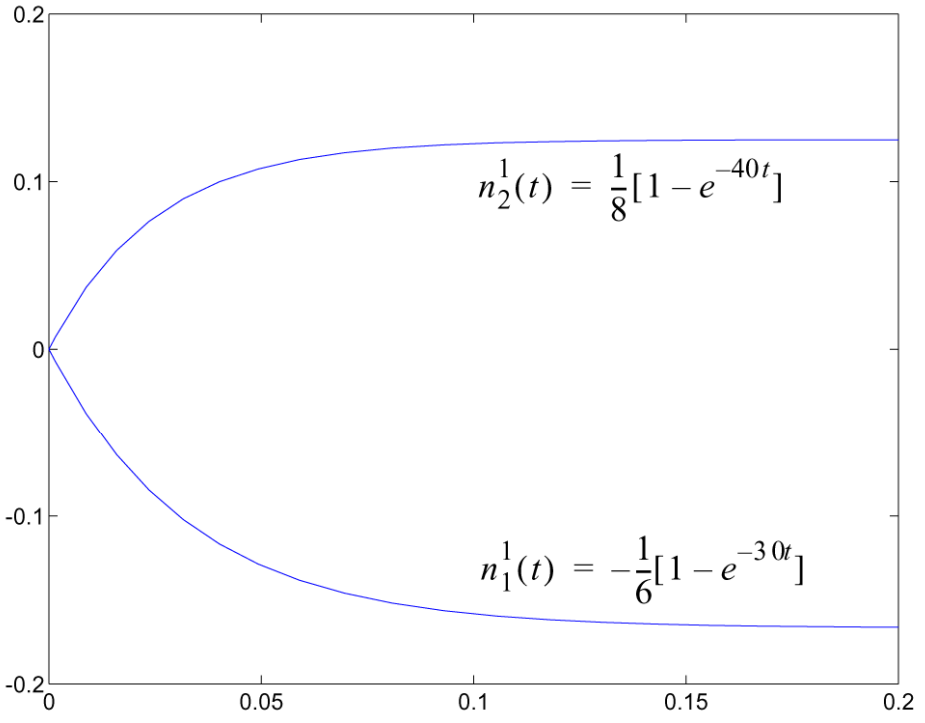
فرض می‌کنیم لایه‌ی ۲ فعال باشد و نرون ۲–ام آن رقابت را برده باشد:

$$\left.\begin{aligned}(0.1)\frac{dn_1^1}{dt} &= -n_1^1 + (1-n_1^1)\{p_1 + w_{1,2}^{2:1}\} - (n_1^1 + 1.5) \\ &= -n_1^1 + (1-n_1^1)\{0+1\} - (n_1^1 + 1.5) = -3n_1^1 - 0.5\end{aligned}\right\} \quad \frac{dn_1^1}{dt} = -30n_1^1 - 5$$

معادلات عملیاتی لایه‌ی ۱ :

$$\left.\begin{aligned}(0.1)\frac{dn_2^1}{dt} &= -n_2^1 + (1-n_2^1)\{p_2 + w_{2,2}^{2:1}\} - (n_2^1 + 1.5) \\ &= -n_2^1 + (1-n_2^1)\{1+1\} - (n_2^1 + 1.5) = -4n_2^1 + 0.5\end{aligned}\right\} \quad \frac{dn_2^1}{dt} = -40n_2^1 + 5$$

Example Response



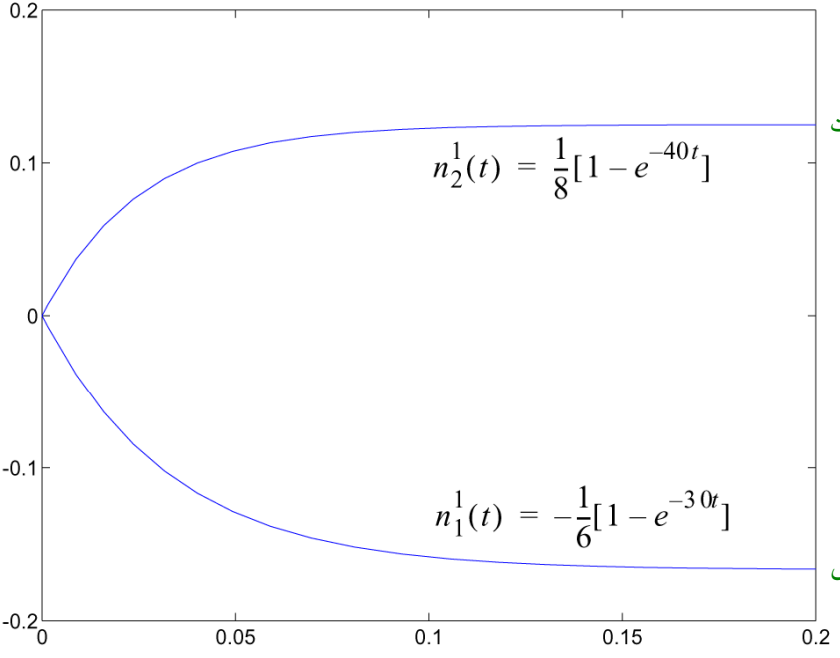
$$\mathbf{p} \cap \mathbf{w}_2^{2:1} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \cap \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \mathbf{a}^1$$

نظریه‌ی تشدید وفقی

لایه‌ی ۱ : مثال : پاسخ

EXAMPLE RESPONSE

با فرض شرایط آغازین صفر



همگرا به یک مقدار مثبت

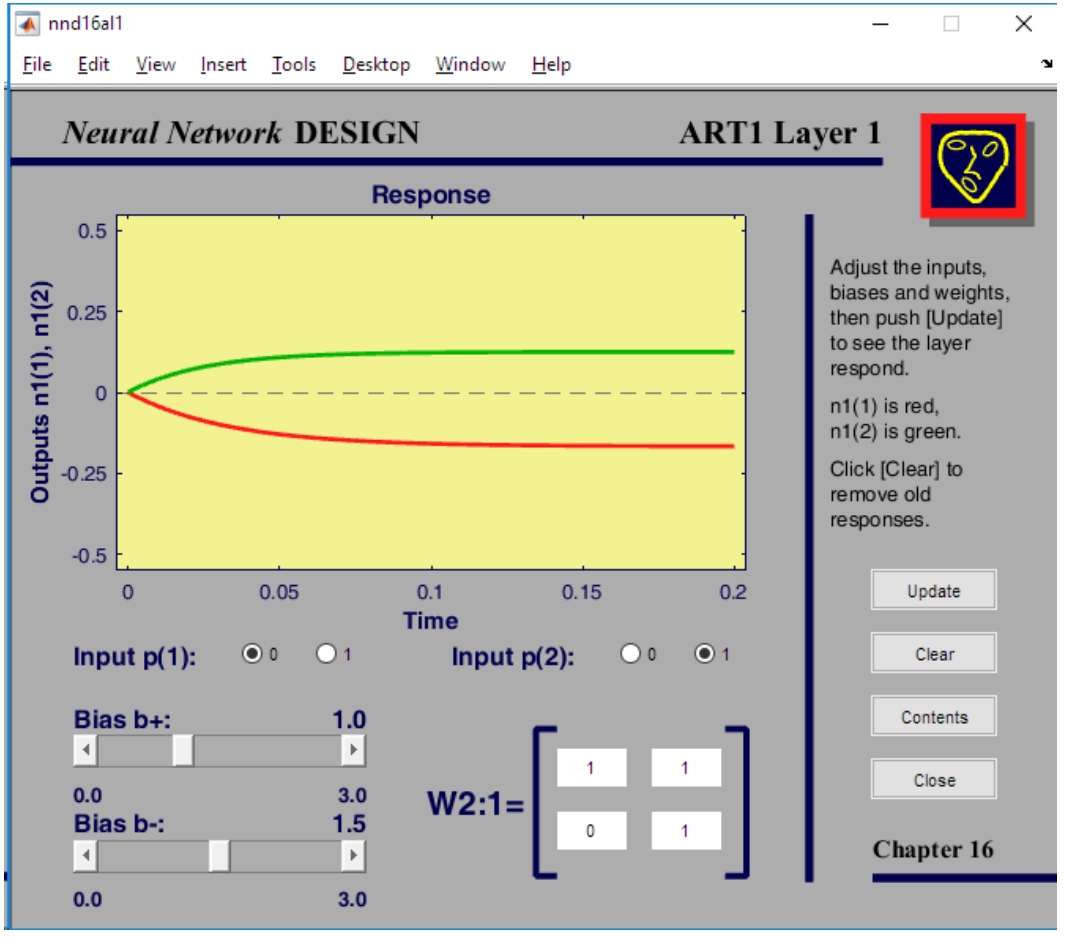
$$a_1^1(t) \to 0$$

همگرا به یک مقدار منفی

$$a_2^1(t) \to 1$$

$$\mathbf{p} \cap \mathbf{w}_2^{2:1} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \cap \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \mathbf{a}^1$$

این نتایج با تحلیل ما از حالت ماندگار همخوانی دارند :



MATLAB

```
>> nnd16al1
```

شبکه‌های عصبی مصنوعی

نظریه‌ی تشدید وفقی

۳

لایه‌ی ۲

نظریه‌ی تشدید وفقی

لایه‌ی ۲

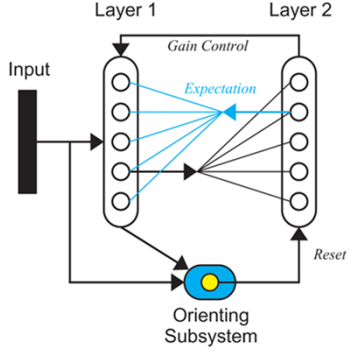
LAYER 2

هدف اصلی لایه‌ی ۲، ارتقای کنتراست در الگوی خروجی آن است.

در این پیاده‌سازی، ارتقای کنتراست از طریق یک رقابت «برنده همه را می‌خورد» انجام می‌شود.
⇐ نرون دارای بیشترین ورودی دارای خروجی غیر صفر است.

در لایه‌ی ۲، انتگرال‌گیر دارای قابلیت بازنشانی است:
خروجی‌هایی که بازنشانی می‌شوند، برای یک دوره‌ی طولانی در صفر باقی می‌مانند.
(دوره‌ی طولانی == مدتی که یک انطباق کافی رخ بدهد و وزن‌ها به‌روزرسانی شوند.)

پیشنهاد کارپنتر و گراسبرگ برای پیاده‌سازی مکانیسم بازنشانی: استفاده از یک درگاه دوقطبی (← مدل پیچیده‌تر در ART3)



لایه‌ی ۲ شبکه‌ی ART1 بسیار شبیه به لایه‌ی ۲ شبکه‌ی گراسبرگ است.

تفاوت‌ها:
- استفاده از انتگرال‌گیر با قابلیت RESET در شبکه‌ی ART1
- استفاده از دو تابع انتقال

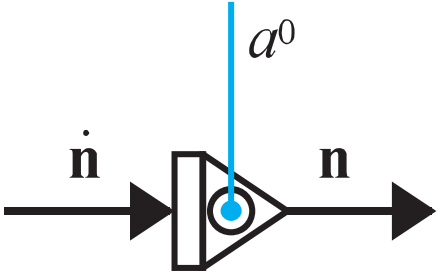
نظریه‌ی تشدید وفقی

انتگرال‌گیر با قابلیت بازنشانی

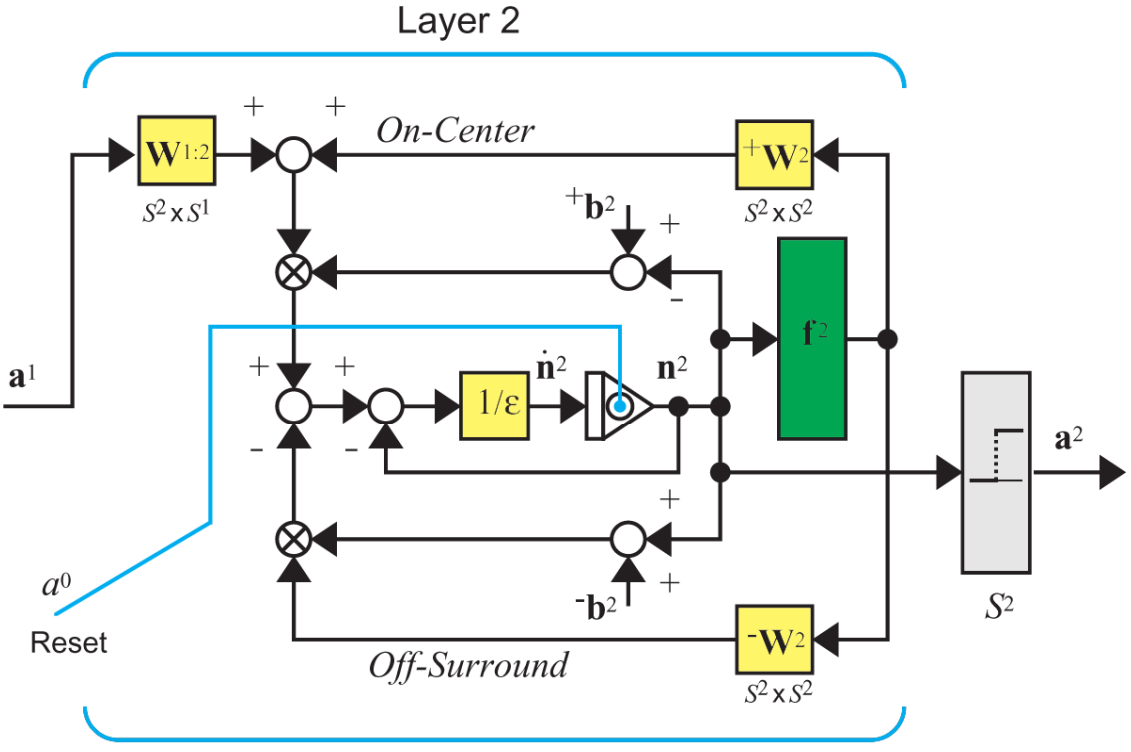
INTEGRATOR WITH RESET CAPABILITY

انتگرال‌گیر با قابلیت بازنشانی

در این نوع انتگرال‌گیر، هر خروجی مثبت به 0 بازنشانی (reset) می‌شود
اگر سیگنال a^0 مثبت باشد.



Layer 2

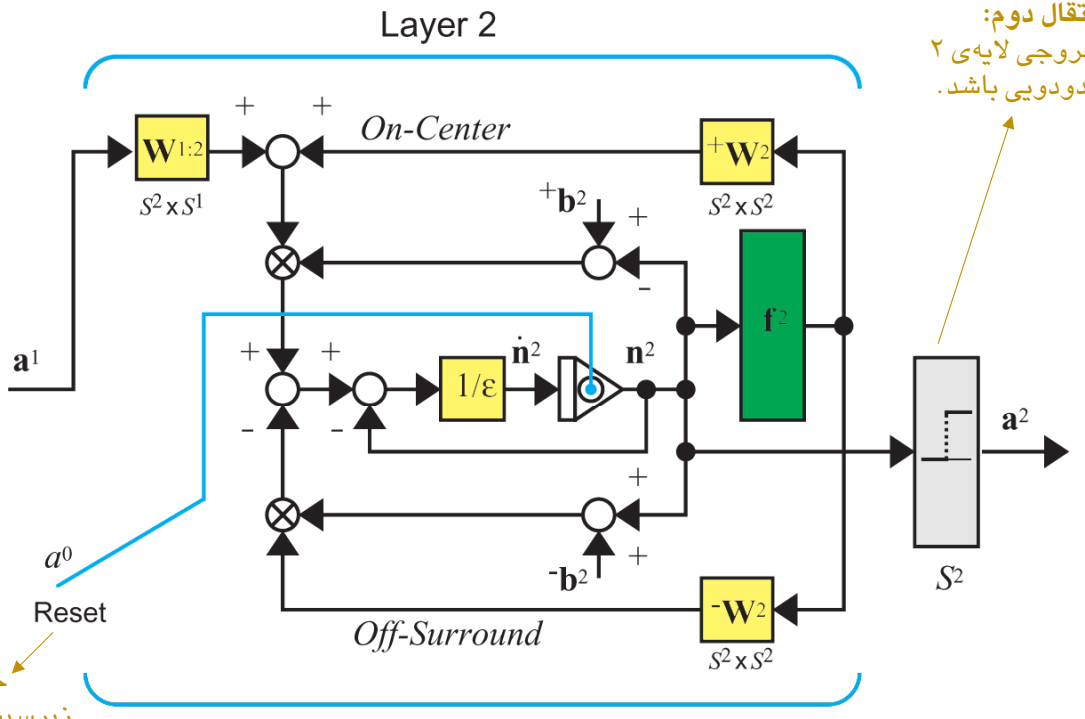


$$\varepsilon \, d\mathbf{n}^2/dt = -\mathbf{n}^2 + (^{+}\mathbf{b}^2 - \mathbf{n}^2)\{[^{+}\mathbf{W}^2]\mathbf{f}^2(\mathbf{n}^2) + \mathbf{W}^{1:2}\mathbf{a}^1\}$$
$$- (\mathbf{n}^2 + {}^{-}\mathbf{b}^2)[^{-}\mathbf{W}^2]\mathbf{f}^2(\mathbf{n}^2)$$

نظریه‌ی تشدید وفقی

لایه‌ی ۲

LAYER 2



تابع انتقال دوم:
می‌خواهیم خروجی لایه‌ی ۲ یک سیگنال دودویی باشد.

خروجی زیرسیستم گرادهنده

$$\varepsilon\, d\mathbf{n}^2/dt = -\mathbf{n}^2 + (^{+}\mathbf{b}^2 - \mathbf{n}^2)\{[^{+}\mathbf{W}^2]\mathbf{f}^2(\mathbf{n}^2) + \mathbf{W}^{1:2}\mathbf{a}^1\} - (\mathbf{n}^2 + {}^{-}\mathbf{b}^2)[^{-}\mathbf{W}^2]\mathbf{f}^2(\mathbf{n}^2)$$

Layer 2 Operation

Shunting Model

$$\varepsilon \frac{d\mathbf{n}^2(t)}{dt} = -\mathbf{n}^2(t) + (^{+}\mathbf{b}^2 - \mathbf{n}^2(t))\{[^{+}\mathbf{W}^2]\mathbf{f}^2(\mathbf{n}^2(t)) + \mathbf{W}^{1:2}\mathbf{a}^1\} - (\mathbf{n}^2(t) + {}^{-}\mathbf{b}^2)[^{-}\mathbf{W}^2]\mathbf{f}^2(\mathbf{n}^2(t))$$

On-Center Feedback

Adaptive Instars

Excitatory Input

Off-Surround Feedback

Inhibitory Input

نظریه‌ی تشدید وفقی

عملیات لایه‌ی ۲

LAYER 2 OPERATION

Shunting Model

$$\varepsilon \frac{d\mathbf{n}^2(t)}{dt} = -\mathbf{n}^2(t) + (^{+}\mathbf{b}^2 - \mathbf{n}^2(t))\{[^{+}\mathbf{W}^2]\mathbf{f}^2(\mathbf{n}^2(t)) + \mathbf{W}^{1:2}\mathbf{a}^1\} - (\mathbf{n}^2(t) + ^{-}\mathbf{b}^2)[^{-}\mathbf{W}^2]\mathbf{f}^2(\mathbf{n}^2(t))$$

اتصالات فیدبک (مشابه لایه ۱ و ۲ گراسبرگ)

On-Center Feedback

وزن‌های وفقی (معادل با شبکه‌ی کوهونن)

Adaptive Instars

Excitatory Input

ورودی تهییج‌گر

اتصالات فیدبک (مشابه لایه ۱ و ۲ گراسبرگ)

Off-Surround Feedback

Inhibitory Input

ورودی بازدارنده

سطرهای ماتریس $\mathbf{W}^{1:2}$ پس از آموزش بازنمایی‌کننده‌ی الگوهای پروتوتایپ می‌باشند.

Layer 2 Example

$$\varepsilon = 0.1 \qquad {}^{+}\mathbf{b}^2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad {}^{-}\mathbf{b}^2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad \mathbf{W}^{1:2} = \begin{bmatrix} ({}_1\mathbf{w}^{1:2})^T \\ ({}_2\mathbf{w}^{1:2})^T \end{bmatrix} = \begin{bmatrix} 0.5 & 0.5 \\ 1 & 0 \end{bmatrix}$$

$$f^2(n) = \begin{cases} 10(n)^2, & n \geq 0 \\ 0, & n < 0 \end{cases}$$

(Faster than linear, winner-take-all)

$$(0.1)\frac{dn_1^2(t)}{dt} = -n_1^2(t) + (1 - n_1^2(t))\left\{ f^2(n_1^2(t)) + ({}_1\mathbf{w}^{1:2})^T \mathbf{a}^1 \right\} - (n_1^2(t) + 1) f^2(n_2^2(t))$$

$$(0.1)\frac{dn_2^2(t)}{dt} = -n_2^2(t) + (1 - n_2^2(t))\left\{ f^2(n_2^2(t)) + ({}_2\mathbf{w}^{1:2})^T \mathbf{a}^1 \right\} - (n_2^2(t) + 1) f^2(n_1^2(t)) .$$

نظریه‌ی تشدید وفقی

لایه‌ی ۲ : مثال

LAYER 2 EXAMPLE

یک لایه با دو نرون و پارامترهای زیر را در نظر می‌گیریم:

$$\varepsilon = 0.1 \qquad {}^{+}\mathbf{b}^2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad {}^{-}\mathbf{b}^2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad \mathbf{W}^{1:2} = \begin{bmatrix} ({}_1\mathbf{w}^{1:2})^T \\ ({}_2\mathbf{w}^{1:2})^T \end{bmatrix} = \begin{bmatrix} 0.5 & 0.5 \\ 1 & 0 \end{bmatrix}$$

تابع انتقال سریع‌تر از خطی
(⇐ برنده همه را می‌خورد.)

$$f^2(n) = \begin{cases} 10(n)^2, & n \geq 0 \\ 0, & n < 0 \end{cases} \qquad \text{(Faster than linear, winner-take-all)}$$

معادلات عملیاتی لایه‌ی ۲ :

$$(0.1)\frac{dn_1^2(t)}{dt} = -n_1^2(t) + (1 - n_1^2(t))\left\{ f^2(n_1^2(t)) + ({}_1\mathbf{w}^{1:2})^T \mathbf{a}^1 \right\} - (n_1^2(t) + 1) f^2(n_2^2(t))$$

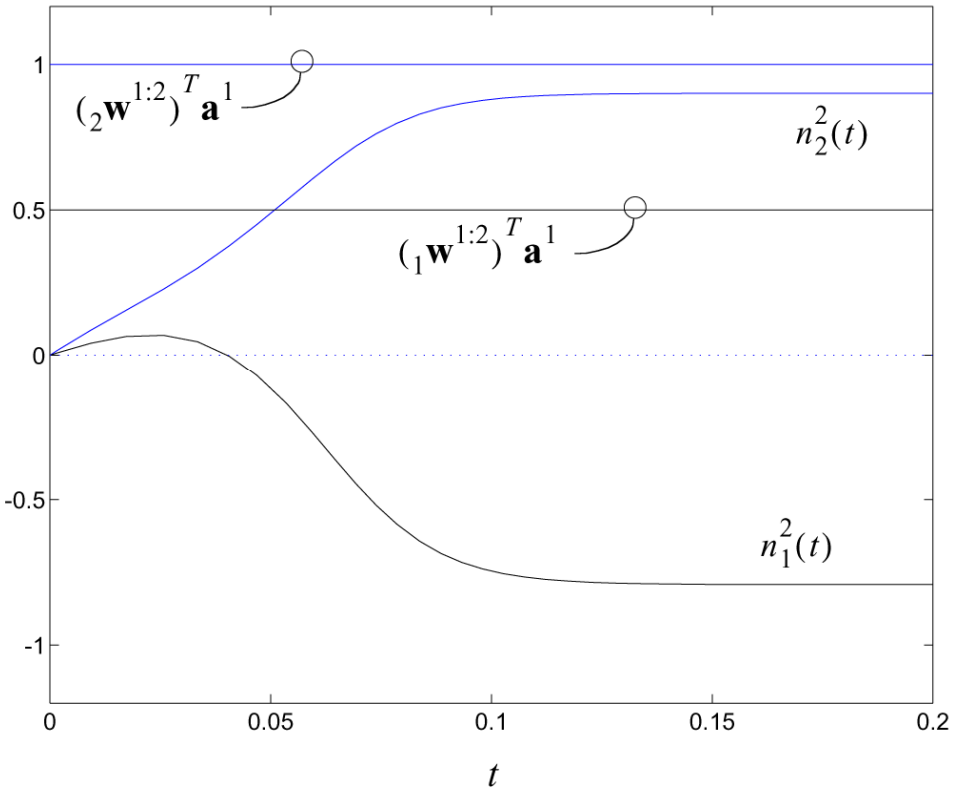
مشابه معادلات لایه ۲ شبکه گراسبرگ، به‌استثنای اینکه $-b^2 = 1 \Leftarrow {}^{-}b^2 = 1$ و $n_2{}^2$ و $n_1{}^2$ در بازه‌ی [1+ , 1-] قرار می‌گیرند.

$$(0.1)\frac{\mathrm{d}n_2^2(t)}{\mathrm{d}t} = -n_2^2(t) + (1 - n_2^2(t))\left\{ f^2(n_2^2(t)) + ({}_2\mathbf{w}^{1:2})^T \mathbf{a}^1 \right\} - (n_2^2(t) + 1) f^2(n_1^2(t)) .$$

ورودی‌های لایه‌ی ۲ : ضرب داخلی الگوهای پروتوتایپ (سطرهای ماتریس $\mathbf{W}^{1:2}$) با خروجی لایه‌ی ۱ ($\mathbf{a}^1$)

Example Response

$$\mathbf{a}^1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$



$$\mathbf{a}^2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

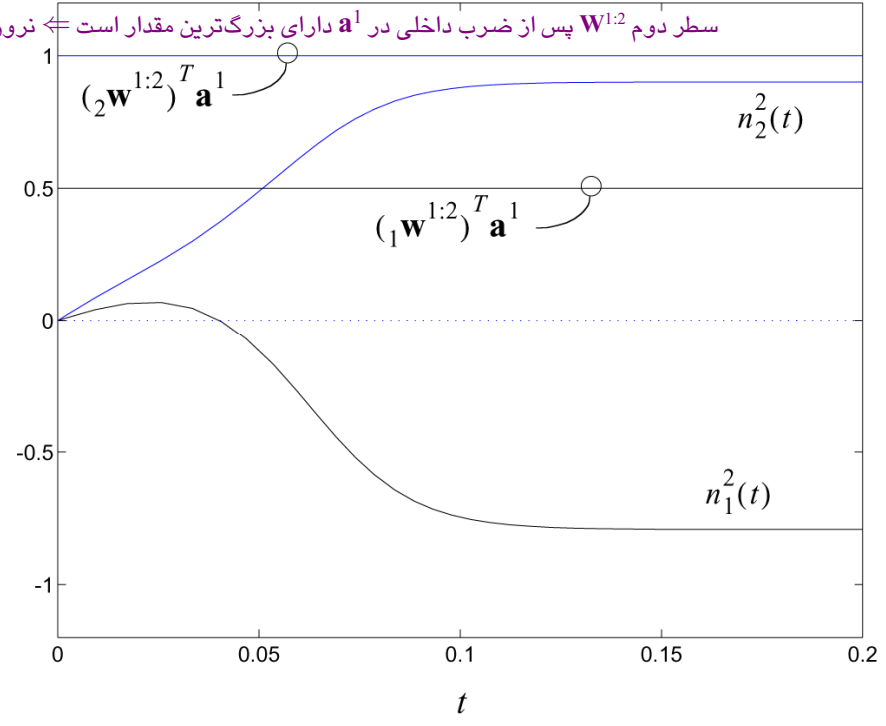
نظریه‌ی تشدید وفقی

لایه‌ی ۲ : مثال : پاسخ

EXAMPLE RESPONSE

پس از انجام رقابت، فقط یک نرون دارای خروجی 1 خواهد بود و خروجی سایر نرون‌ها صفر است (تابع انتقال hardlim).

سطر دوم $\mathbf{W}^{1:2}$ پس از ضرب داخلی در $\mathbf{a}^1$ دارای بزرگ‌ترین مقدار است ⇐ نرون ۲ رقابت را می‌برد.



$$\mathbf{a}^1 = \begin{bmatrix}1\\0\end{bmatrix}$$

مقدار مثبت در حالت ماندگار

$$\mathbf{a}^2 = \begin{bmatrix}0\\1\end{bmatrix}$$

مقدار منفی در حالت ماندگار

Layer 2 Summary

$$a_i^2 = \begin{cases} 1, & \text{if}(({}_i\mathbf{w}^{1:2})^T \mathbf{a}^1 = max[({}_j\mathbf{w}^{1:2})^T \mathbf{a}^1]) \\ 0, & \text{otherwise} \end{cases}$$

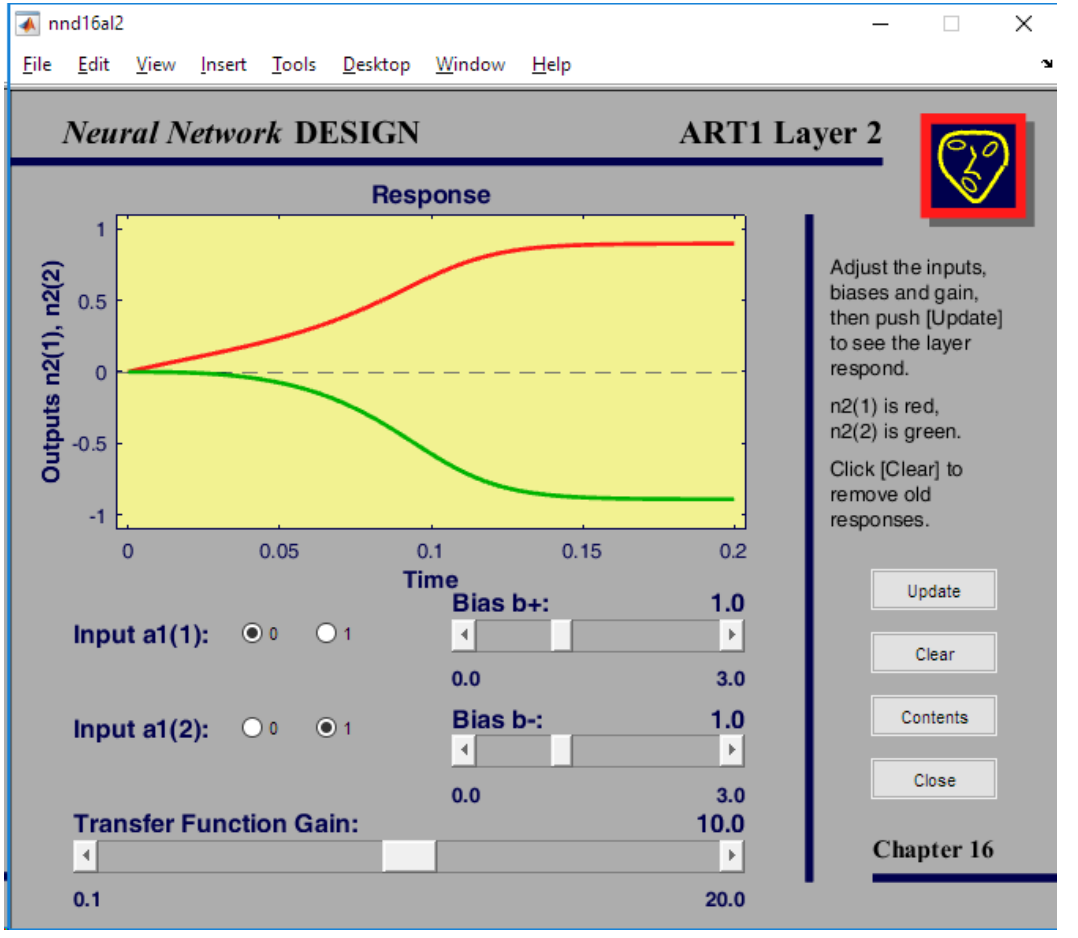
نظریه‌ی تشدید وفقی

خلاصه‌ی لایه‌ی ۲

LAYER 2 SUMMARY

$$a_i^2 = \begin{cases} 1, & \text{if}((_i\mathbf{w}^{1:2})^T \mathbf{a}^1 = max[(_j\mathbf{w}^{1:2})^T \mathbf{a}^1]) \\ 0, & \text{otherwise} \end{cases}$$

خروجی نرون برنده 1 و سایر نرون‌ها 0


nnd16al2
File Edit View Insert Tools Desktop Window Help
Neural Network DESIGN
ART1 Layer 2
Response
Outputs n2(1), n2(2)
1
0.5
0
-0.5
-1
0
0.05
0.1
0.15
0.2
Time
Adjust the inputs, biases and gain, then push [Update] to see the layer respond.
n2(1) is red, n2(2) is green.
Click [Clear] to remove old responses.
Update
Clear
Contents
Close
Chapter 16
Input a1(1): 0 1
Bias b+: 1.0
0.0 3.0
Input a1(2): 0 1
Bias b-: 1.0
0.0 3.0
Transfer Function Gain: 10.0
0.1 20.0

```
>> nnd16al2
```

شبکه‌های عصبی مصنوعی

نظریه‌ی تشدید وفقی

۴

زیرسیستم گرادهنده

نظریه‌ی تشدید وفقی

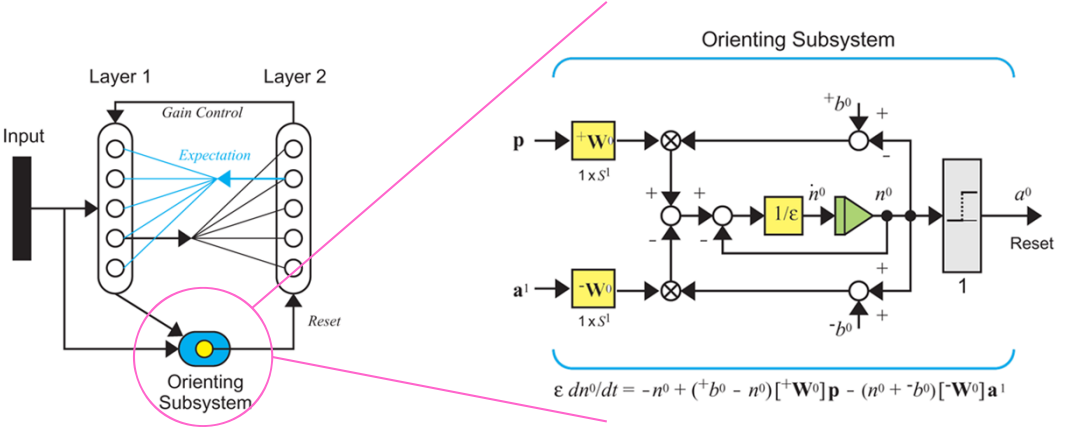
زیرسیستم گرادهنده

ORIENTING SUBSYSTEM

هدف زیرسیستم گرادهنده: تشخیص وجود انطباق کافی بین انتظار L2-L1 ($\mathbf{a}^1$) و الگوی ورودی ($\mathbf{p}$).

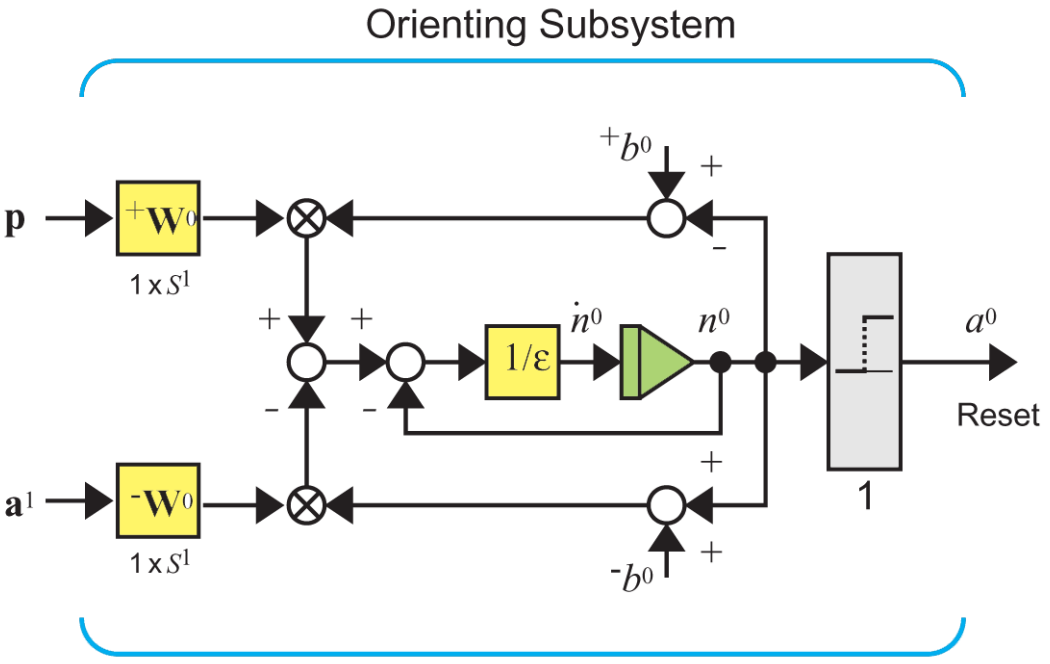
(زیرسیستم گرادهنده یک جزء کلیدی در معماری ART است)

اگر انطباق کافی وجود نداشته باشد،
زیرسیستم گرادهنده باید یک سیگنال بازنشانی reset به لایه‌ی ۲ ارسال کند.
این سیگنال بازنشانی منجر به یک بازدارندگی طولانی مدت بر روی نرون برنده می‌شود
و در نتیجه سایر نرون‌ها را قادر به پیروزی در رقابت می‌کند.



Prepared by Kazim Fouladi | Fall 2017 | 2nd Edition
کاظم فولادی

Orienting Subsystem



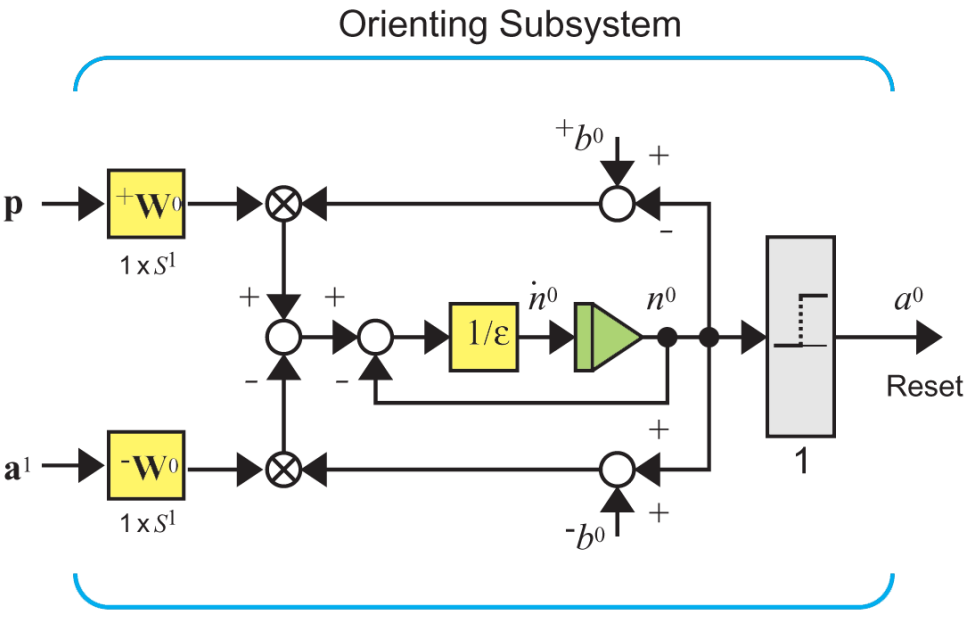
$$\varepsilon\, dn^0/dt = -n^0 + ({}^{+}b^0 - n^0)[{}^{+}\mathbf{W}^0]\mathbf{p} - (n^0 + {}^{-}b^0)[{}^{-}\mathbf{W}^0]\mathbf{a}^1$$

Purpose: Determine if there is a sufficient match between the L2-L1 expectation ($\mathbf{a}^1$) and the input pattern ($\mathbf{p}$).

نظریه‌ی تشدید وفقی

زیرسیستم گرادهنده

ORIENTING SUBSYSTEM



$$\varepsilon\, dn^0/dt = -n^0 + (^{+}b^0 - n^0)[^{+}\mathbf{W}^0]\mathbf{p} - (n^0 + {}^{-}b^0)[^{-}\mathbf{W}^0]\mathbf{a}^1$$

Purpose: Determine if there is a sufficient match between the L2-L1 expectation ($\mathbf{a}^1$) and the input pattern ($\mathbf{p}$).

Orienting Subsystem Operation

$$\varepsilon \frac{dn^0(t)}{dt} = -n^0(t) + ({}^{+}b^0 - n^0(t))\{{}^{+}\mathbf{W}^0\mathbf{p}\} - (n^0(t) + {}^{-}b^0)\{{}^{-}\mathbf{W}^0\mathbf{a}^1\}$$

Excitatory Input

$${}^{+}\mathbf{W}^0\mathbf{p} = [\alpha\ \alpha\ \ldots\ \alpha]\mathbf{p} = \alpha \sum_{j=1}^{S^1} p_j = \alpha \|\mathbf{p}\|^2$$

Inhibitory Input

$${}^{-}\mathbf{W}^0\mathbf{a}^1 = [\beta\ \beta\ \ldots\ \beta]\mathbf{a}^1 = \beta \sum_{j=1}^{S^1} a_j^1(t) = \beta \|\mathbf{a}^1\|^2$$

When the excitatory input is larger than the inhibitory input, the Orienting Subsystem will be driven on.

نظریه‌ی تشدید وفقی

عملیات زیرسیستم گرادهنده

ORIENTING SUBSYSTEM OPERATION

یک مدل شانتینگ:

$$\varepsilon \frac{dn^0(t)}{dt} = -n^0(t) + ({}^{+}b^0 - n^0(t))\{{}^{+}\mathbf{W}^0\mathbf{p}\} - (n^0(t) + {}^{-}b^0)\{{}^{-}\mathbf{W}^0\mathbf{a}^1\}$$

ورودی تهییج‌گر Excitatory Input

$${}^{+}\mathbf{W}^0\mathbf{p} = [\alpha\ \alpha\ \dots\ \alpha]\mathbf{p} = \alpha\sum_{j=1}^{S^1} p_j = \alpha\|\mathbf{p}\|^2$$

چون بردار $\mathbf{p}$ دودویی است.

ورودی بازدارنده Inhibitory Input

$${}^{-}\mathbf{W}^0\mathbf{a}^1 = [\beta\ \beta\ \dots\ \beta]\mathbf{a}^1 = \beta\sum_{j=1}^{S^1} a_j^1(t) = \beta\|\mathbf{a}^1\|^2$$

When the excitatory input is larger than the inhibitory input, the Orienting Subsystem will be driven on.

هرگاه ورودی تهییج‌گر بزرگ‌تر از ورودی بازدارنده باشد، زیرسیستم گرادهنده فعال می‌شود.

Steady State Operation

$$0 = -n^0 + ({}^{+}b^0 - n^0)\{\alpha\|\mathbf{p}\|^2\} - (n^0 + {}^{-}b^0)\left\{\beta\|\mathbf{a}^1\|^2\right\}$$

$$= -(1 + \alpha\|\mathbf{p}\|^2 + \beta\|\mathbf{a}^1\|^2)n^0 + {}^{+}b^0(\alpha\|\mathbf{p}\|^2) - {}^{-}b^0(\beta\|\mathbf{a}^1\|^2)$$

$$n^0 = \frac{{}^{+}b^0(\alpha\|\mathbf{p}\|^2) - {}^{-}b^0(\beta\|\mathbf{a}^1\|^2)}{(1 + \alpha\|\mathbf{p}\|^2 + \beta\|\mathbf{a}^1\|^2)}$$

Let ${}^{+}b^0 = {}^{-}b^0 = 1$

$n^0 > 0$ if $\dfrac{\|\mathbf{a}^1\|^2}{\|\mathbf{p}\|^2} < \dfrac{\alpha}{\beta} = \rho$

RESET

ρ → Vigilance

Since $\mathbf{a}^1 = \mathbf{p} \cap \mathbf{w}_j^{2:1}$, a reset will occur when there is enough of a mismatch between $\mathbf{p}$ and $\mathbf{w}_j^{2:1}$.

نظریه‌ی تشدید وفقی

عملیات زیرسیستم گرادهنده در حالت ماندگار

STEADY STATE OPERATION

(مشتق در حالت ماندگار صفر است)

$$0 = -n^0 + ({}^{+}b^0 - n^0)\{\alpha\|\mathbf{p}\|^2\} - (n^0 + {}^{-}b^0)\left\{\beta\left\|\mathbf{a}^1\right\|^2\right\}$$

$$= -(1 + \alpha\|\mathbf{p}\|^2 + \beta\left\|\mathbf{a}^1\right\|^2)n^0 + {}^{+}b^0(\alpha\|\mathbf{p}\|^2) - {}^{-}b^0(\beta\left\|\mathbf{a}^1\right\|^2)$$

$$n^0 = \frac{{}^{+}b^0(\alpha\|\mathbf{p}\|^2) - {}^{-}b^0(\beta\left\|\mathbf{a}^1\right\|^2)}{(1 + \alpha\|\mathbf{p}\|^2 + \beta\left\|\mathbf{a}^1\right\|^2)}$$

فرض می‌کنیم: $\text{Let} \quad {}^{+}b^0 = {}^{-}b^0 = 1$

در صورت فعال بودن لایه‌ی ۲، $\mathbf{a}^1 = \mathbf{p} \cap \mathbf{w}_j^{2:1}$ خواهد بود.
پس $\|\mathbf{p}\|^2 \geq \|\mathbf{a}^1\|^2$ و زیرسیستم گرادهنده در صورت عدم تطابق کافی بین $\mathbf{p}$ و $\mathbf{w}_j^{2:1}$ منجر به بازنشانی می‌شود.

$$n^0 > 0 \quad \text{if} \quad \frac{\left\|\mathbf{a}^1\right\|^2}{\|\mathbf{p}\|^2} < \frac{\alpha}{\beta} = \rho$$

RESET

میزان عدم تطابق برای انجام بازنشانی، توسط پارامتر هوشیاری ρ تعیین می‌شود.

Vigilance

پارامتر هوشیاری

Since $\mathbf{a}^1 = \mathbf{p} \cap \mathbf{w}_j^{2:1}$, a reset will occur when there is enough of a mismatch between $\mathbf{p}$ and $\mathbf{w}_j^{2:1}$.

نظریه‌ی تشدید وفقی

زیرسیستم گرادهنده: پارامتر هوشیاری

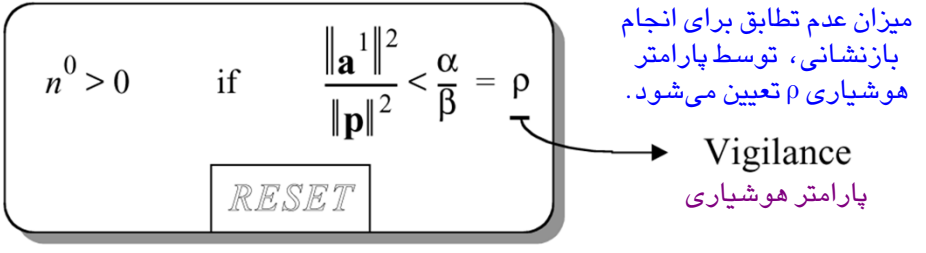
VIGILANCE PARAMETER

پارامتر هوشیاری

$$0 < \rho < 1$$

- اگر نزدیک به 1 باشد، عملیات reset در صورتی که $\mathbf{a}^1$ به $\mathbf{p}$ نزدیک نباشد رخ می‌دهد.
- اگر نزدیک به 0 باشد، نیازی به نزدیک بودن $\mathbf{a}^1$ به $\mathbf{p}$ برای جلوگیری از reset وجود ندارد.

* پارامتر هوشیاری میزان سخت‌گیری روال خوشه‌بندی توسط بردارهای پروتوتایپ را مشخص می‌کند.

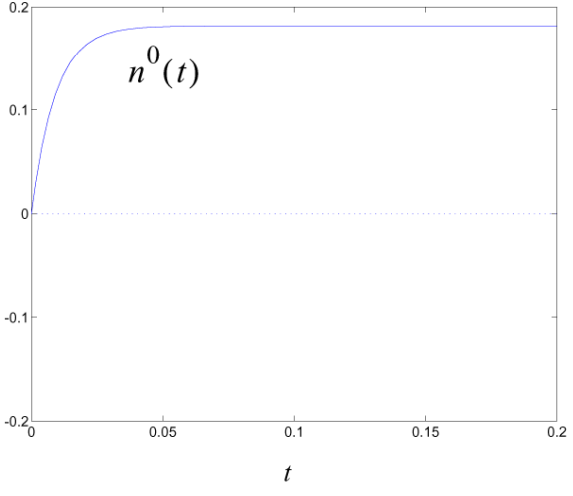


Orienting Subsystem Example

$$\varepsilon = 0.1,\ \alpha = 3,\ \beta = 4\ (\rho = 0.75) \qquad \mathbf{p} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad \mathbf{a}^1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$(0.1)\frac{dn^0(t)}{dt} = -n^0(t) + (1 - n^0(t))\{3(p_1 + p_2)\} - (n^0(t) + 1)\{4(a_1^1 + a_2^1)\}$$

$$\frac{dn^0(t)}{dt} = -110n^0(t) + 20$$



نظریه‌ی تشدید وفقی

زیرسیستم گرادهنده: مثال

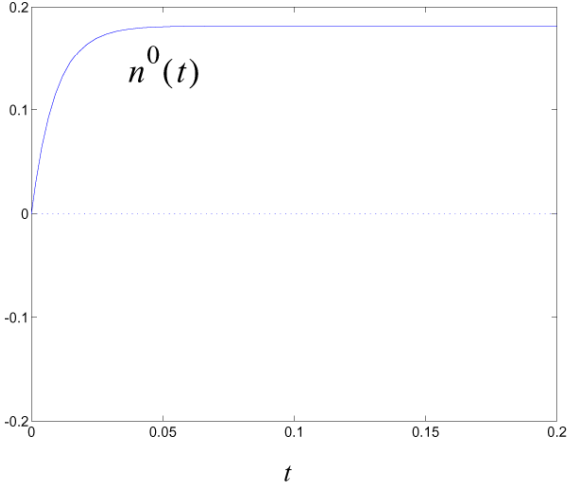
ORIENTING SUBSYSTEM EXAMPLE

$$\varepsilon = 0.1,\ \alpha = 3,\ \beta = 4\ (\rho = 0.75) \qquad \mathbf{p} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad \mathbf{a}^1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

معادله‌ی عملیاتی:

$$(0.1)\frac{dn^0(t)}{dt} = -n^0(t) + (1 - n^0(t))\{3(p_1 + p_2)\} - (n^0(t) + 1)\{4(a_1^1 + a_2^1)\}$$

$$\frac{dn^0(t)}{dt} = -110n^0(t) + 20$$



سیگنال reset به لایه‌ی ۲ ارسال می‌شود، زیرا $n^0(t)$ مثبت است.

در این مثال: از آنجا که پارامتر هوشیاری برابر 0.75 است و $\mathbf{p}$ تنها دارای دو عضو است، reset تنها در صورتی رخ می‌دهد که $\mathbf{p}$ و $\mathbf{a}^1$ یکسان نباشند.

در این مثال: اگر پارامتر هوشیاری برابر 0.25 بود هیچ بازنشانی رخ نمی‌داد، زیرا $\|\mathbf{a}^1\|^2/\|\mathbf{p}\|^2 = 1/2$

Orienting Subsystem Summary

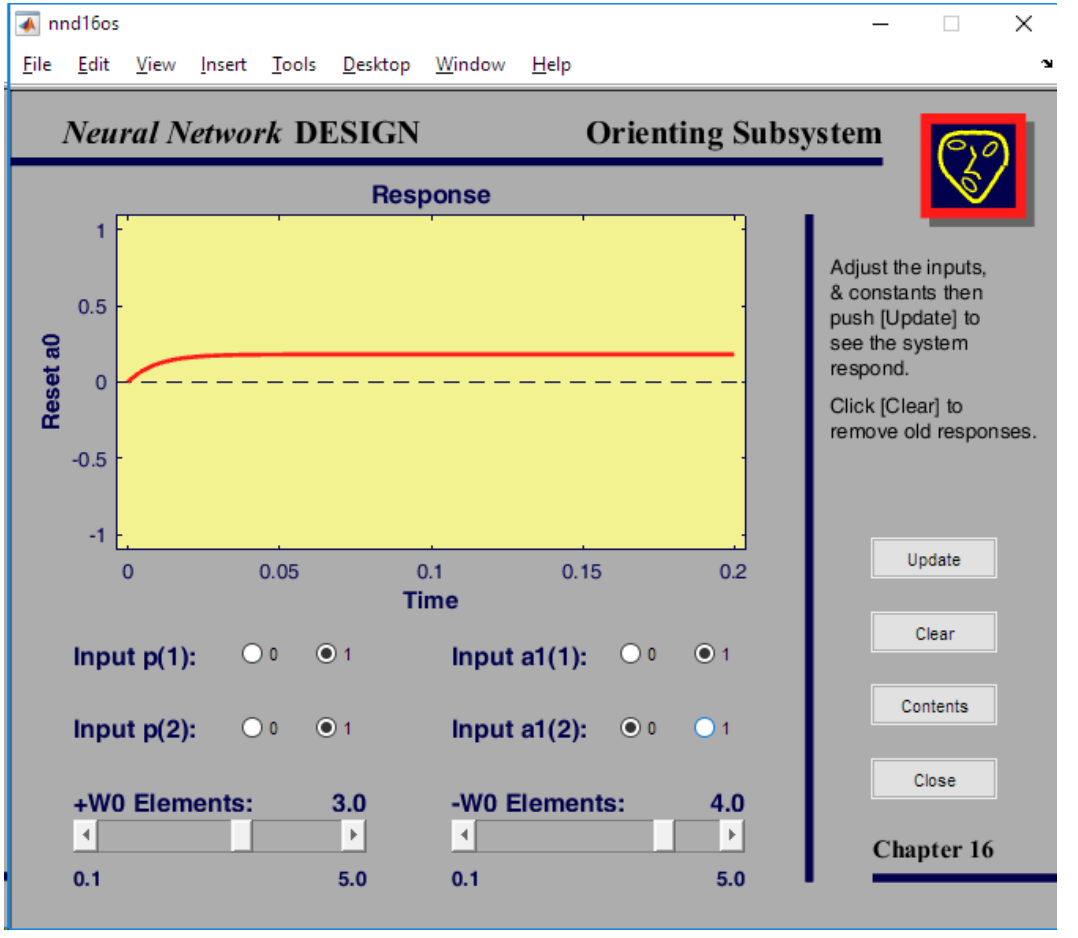
$$a^0 = \begin{cases} 1, & \text{if}[\|\mathbf{a}^1\|^2 / \|\mathbf{p}\|^2 < \rho] \\ 0, & \text{otherwise} \end{cases}$$

نظریه‌ی تشدید وفقی

زیرسیستم گرادهنده: خلاصه

ORIENTING SUBSYSTEM SUMMARY

$$a^0 = \begin{cases} 1, & \text{if}[\|\mathbf{a}^1\|^2 / \|\mathbf{p}\|^2 < \rho] \\ 0, & \text{otherwise} \end{cases}$$


nnd16os
File Edit View Insert Tools Desktop Window Help
Neural Network DESIGN
Orienting Subsystem
Response
Reset a0
1
0.5
0
-0.5
-1
0
0.05
0.1
0.15
0.2
Time
Adjust the inputs, & constants then push [Update] to see the system respond.
Click [Clear] to remove old responses.
Update
Clear
Contents
Close
Chapter 16
Input p(1): 0 1
Input p(2): 0 1
Input a1(1): 0 1
Input a1(2): 0 1
+W0 Elements: 3.0
0.1 5.0
-W0 Elements: 4.0
0.1 5.0

```
>> nnd16os
```

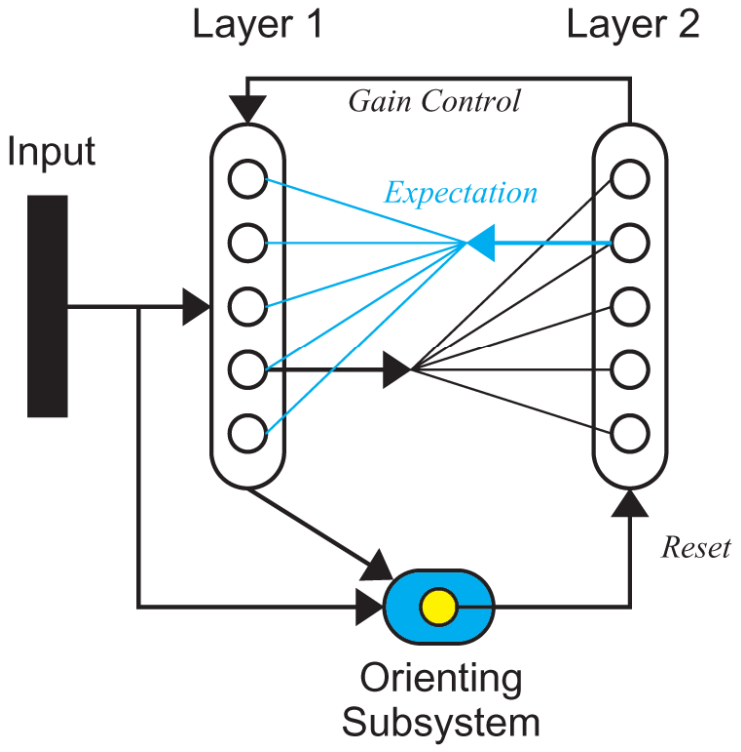
شبکه‌های عصبی مصنوعی

نظریه‌ی تشدید وفقی

۵

قوانین یادگیری

Learning Laws: L1-L2 and L2-L1



The ART1 network has two separate learning laws: one for the L1-L2 connections (instars) and one for the L2-L1 connections (outstars).

Both sets of connections are updated at the same time - when the input and the expectation have an adequate match.

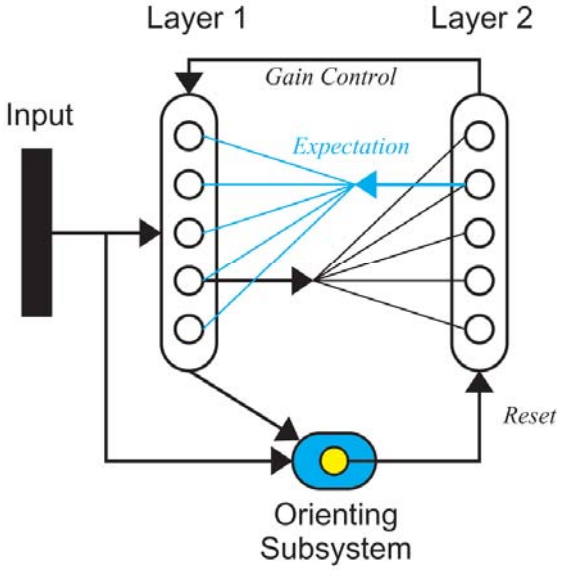
The process of matching, and subsequent adaptation is referred to as resonance.

نظریه‌ی تشدید وفقی

قوانین یادگیری: L1-L2 و L2-L1

LEARNING LAWS: L1-L2 AND L2-L1

شبکه‌ی ART1 دارای دو قاعده‌ی یادگیری مجزا است:
یکی برای اتصالات L1-L2 و دیگری برای اتصالات L2-L1



یادگیری اتصالات L1-L2 (instars)	برای یادگیری بازشناسی مجموعه‌ای از پروتوتایپ‌ها
یادگیری اتصالات L2-L1 (outstars)	برای بازتولید (و به‌یادآوری) مجموعه‌ای از پروتوتایپ‌ها

Both sets of connections are updated at the same time - when the input and the expectation have an adequate match.

The process of matching, and subsequent adaptation is referred to as resonance.

هر دو مجموعه از اتصالات به‌طور همزمان به‌روزرسانی می‌شوند: وقتی ورودی و انتظار انطباق کافی دارند.
(تعیین انطباق کافی با زیرسیستم گرادهنده)

از فرآیند انطباق‌سنجی و وفق‌دهی پس از آن، تحت عنوان «رزونانس = **تشدید**» یاد می‌شود.

نظریه‌ی تشدید وفقی

قوانین یادگیری: L1-L2 و L2-L1: معمای زیرمجموعه/ابرمجموعه

SUBSET/SUPERSET DILEMMA

یادگیری اتصالات L1-L2 در شبکه‌ی ART1 بسیار شبیه به یادگیری در شبکه‌های گراسبرگ است.

البته با یک تفاوت مهم:

در شبکه‌ی گراسبرگ الگوهای ورودی در لایه‌ی ۱ نرمال می‌شوند

⇐ همه‌ی الگوهای پروتوتایپ دارای طول یکسانی می‌باشند.

اما در شبکه‌ی ART1 هیچ عملیات نرمال‌سازی در لایه‌ی ۱ انجام نمی‌شود

⇐ اگر یک الگوی پروتوتایپ زیرمجموعه‌ای از یک الگوی پروتوتایپ دیگر باشد، مشکل ایجاد می‌شود:

(معمای زیرمجموعه/ ابرمجموعه)

Subset/Superset Dilemma

Suppose that $\mathbf{W}^{1:2} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$ so the prototypes are ${}_1\mathbf{w}^{1:2} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ ${}_2\mathbf{w}^{1:2} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

We say that ${}_1\mathbf{w}^{1:2}$ is a subset of ${}_2\mathbf{w}^{1:2}$, because ${}_2\mathbf{w}^{1:2}$ has a 1 wherever ${}_1\mathbf{w}^{1:2}$ has a 1.

If the output of layer 1 is $\mathbf{a}^1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ then the input to Layer 2 will be

$$\mathbf{W}^{1:2}\mathbf{a}^1 = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

Both prototype vectors have the same inner product with $\mathbf{a}^1$, even though the first prototype is identical to $\mathbf{a}^1$ and the second prototype is not. This is called the *Subset/Superset* dilemma.

نظریه‌ی تشدید وفقی

قوانین یادگیری: L1-L2 و L2-L1: معمای زیرمجموعه/ابرمجموعه: مثال

SUBSET/SUPERSET DILEMMA

Suppose that $\mathbf{W}^{1:2} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$ so the prototypes are ${}_1\mathbf{w}^{1:2} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ ${}_2\mathbf{w}^{1:2} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

We say that ${}_1\mathbf{w}^{1:2}$ is a subset of ${}_2\mathbf{w}^{1:2}$, because ${}_2\mathbf{w}^{1:2}$ has a 1 wherever ${}_1\mathbf{w}^{1:2}$ has a 1.

If the output of layer 1 is $\mathbf{a}^1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ then the input to Layer 2 will be

$$\mathbf{W}^{1:2}\mathbf{a}^1 = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

Both prototype vectors have the same inner product with $\mathbf{a}^1$, even though the first prototype is identical to $\mathbf{a}^1$ and the second prototype is not. This is called the *Subset/Superset* dilemma.

هر دو بردار پروتوتایپ ضرب داخلی یکسانی با $\mathbf{a}^1$ دارند، حتی اگر پروتوتایپ اول مساوی $\mathbf{a}^1$ باشد و پروتوتایپ دوم نباشد.
به این مشکل «**معمای زیرمجموعه/ ابرمجموعه**» گفته می‌شود.

Subset/Superset Solution

Normalize the prototype patterns.

$$\mathbf{W}^{1:2} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

$$\mathbf{W}^{1:2}\mathbf{a}^{1} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ \frac{2}{3} \end{bmatrix}$$

Now we have the desired result; the first prototype has the largest inner product with the input.

نظریه‌ی تشدید وفقی

قوانین یادگیری: L1-L2 و L2-L1: معمای زیرمجموعه/ابرمجموعه: راه‌حل

SUBSET/SUPERSET SOLUTION

یک راه‌حل نرمال‌سازی بردارهای پروتوتایپ است:

Normalize the prototype patterns.

$$\mathbf{W}^{1:2} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

بنابراین، ورودی لایه‌ی ۲ به صورت زیر می‌شود:

$$\mathbf{W}^{1:2}\mathbf{a}^1 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ \frac{2}{3} \end{bmatrix}$$

Now we have the desired result; the first prototype has the largest inner product with the input.

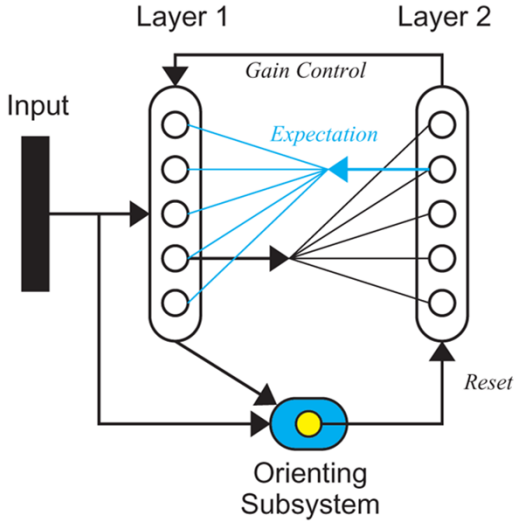
این بار نتیجه‌ی مطلوب حاصل شده است و بردار اولیه در ضرب داخلی با $\mathbf{a}^1$ منجر به مقدار بزرگ‌تری شده است، پس نرون اول از لایه‌ی ۲ فعال می‌شود.

نظریه‌ی تشدید وفقی

مقایسه‌ی نرمال‌سازی در شبکه‌های گراسبرگ و ART1

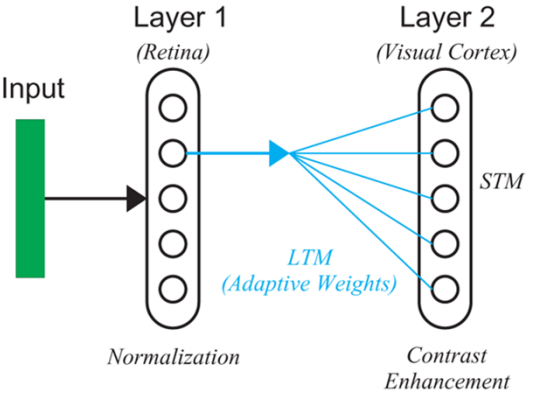
NORMALIZATION

ART



در شبکه‌ی ART1،
عملیات نرمال‌سازی الگوهای پروتوتایپ،
با استفاده از رقابت «مرکز–روشن/ پیرامون–خاموش»
در قانون یادگیری L1-L2 صورت می‌گیرد.

Grossberg



در شبکه‌ی گراسبرگ،
عملیات نرمال‌سازی الگوهای پروتوتایپ،
با نرمال‌سازی الگوهای ورودی در لایه‌ی ۱ انجام می‌شود.

L1-L2 Learning Law

Instar Learning with Competition

$$\frac{d[{}_i\mathbf{w}^{1:2}(t)]}{dt} = a_i^2(t)[\{{}^{+}\mathbf{b} - {}_i\mathbf{w}^{1:2}(t)\}\zeta[{}^{+}\mathbf{W}]\mathbf{a}^1(t) - \{{}_i\mathbf{w}^{1:2}(t) + {}^{-}\mathbf{b}\}[{}^{-}\mathbf{W}]\mathbf{a}^1(t)] ,$$

where

$${}^{+}\mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \quad {}^{-}\mathbf{b} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad {}^{+}\mathbf{W} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \quad {}^{-}\mathbf{W} = \begin{bmatrix} 0 & 1 & \cdots & 1 \\ 1 & 0 & \cdots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \cdots & 0 \end{bmatrix}$$

Upper Limit Bias | Lower Limit Bias | On-Center Connections | Off-Surround Connections

When neuron i of Layer 2 is active, ${}_i\mathbf{w}^{1:2}$ is moved in the direction of $\mathbf{a}^1$. The elements of ${}_i\mathbf{w}^{1:2}$ compete, and therefore ${}_i\mathbf{w}^{1:2}$ is normalized.

نظریه‌ی تشدید وفقی

قانون یادگیری L1-L2

L1-L2 Learning Law

یادگیری Instar با رقابت

Instar Learning with Competition

$$\frac{d[{}_i\mathbf{w}^{1:2}(t)]}{dt} = a_i^2(t)[\{{}^{+}\mathbf{b} - {}_i\mathbf{w}^{1:2}(t)\}\zeta[{}^{+}\mathbf{W}]\mathbf{a}^1(t) - \{{}_i\mathbf{w}^{1:2}(t) + {}^{-}\mathbf{b}\}[{}^{-}\mathbf{W}]\mathbf{a}^1(t)],$$

where

$${}^{+}\mathbf{b} = \begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix} \quad {}^{-}\mathbf{b} = \begin{bmatrix}0\\0\\\vdots\\0\end{bmatrix} \quad {}^{+}\mathbf{W} = \begin{bmatrix}1&0&\cdots&0\\0&1&\cdots&0\\\vdots&\vdots&&\vdots\\0&0&\cdots&1\end{bmatrix} \quad {}^{-}\mathbf{W} = \begin{bmatrix}0&1&\cdots&1\\1&0&\cdots&1\\\vdots&\vdots&&\vdots\\1&1&\cdots&0\end{bmatrix}$$

Upper Limit Bias — Lower Limit Bias — On-Center Connections — Off-Surround Connections

When neuron i of Layer 2 is active, ${}_i\mathbf{w}^{1:2}$ is moved in the direction of $\mathbf{a}^1$. The elements of ${}_i\mathbf{w}^{1:2}$ compete, and therefore ${}_i\mathbf{w}^{1:2}$ is normalized.

وقتی نرون i از لایه‌ی ۲ فعال باشد، ${}_i\mathbf{w}^{1:2}$ در جهت $\mathbf{a}^1$ جابه‌جا می‌شود.
عناصر ${}_i\mathbf{w}^{1:2}$ رقابت می‌کنند و در نتیجه ${}_i\mathbf{w}^{1:2}$ نرمال‌سازی می‌شود.

این معادله بسیار شبیه به مدل شانتینگ با اتصالات ورودی «مرکز-روشن/پیرامون-خاموش» از $\mathbf{a}^1$ است.
بایاس تهییج‌گر در این مورد $\mathbf{1} = {}^{+}\mathbf{b}$ و بایاس بازدارنده $\mathbf{0} = {}^{-}\mathbf{b}$ می‌باشد
و به این‌ترتیب از اینکه اعضای ${}_i\mathbf{w}^{1:2}$ در بازه‌ی $[0,1]$ باشند، اطمینان حاصل می‌شود.

Fast Learning

$$\frac{dw_{i,j}^{1:2}(t)}{dt} = a_i^2(t)\left[(1-w_{i,j}^{1:2}(t))\zeta a_j^1(t) - w_{i,j}^{1:2}(t)\sum_{k\neq j} a_k^1(t)\right]$$

For *fast learning* we assume that the outputs of Layer 1 and Layer 2 remain constant until the weights reach steady state.

Assume that $a^2_i(t) = 1$, and solve for the steady state weight:

$$0 = \left[(1-w_{i,j}^{1:2})\zeta a_j^1 - w_{i,j}^{1:2}\sum_{k\neq j} a_k^1\right]$$

Case I: $a^1_j = 1$

$$0 = (1-w_{i,j}^{1:2})\zeta - w_{i,j}^{1:2}(\|\mathbf{a}^1\|^2 - 1) = -(\zeta + \|\mathbf{a}^1\|^2 - 1)w_{i,j}^{1:2} + \zeta \Rightarrow w_{i,j}^{1:2} = \frac{\zeta}{\zeta + \|\mathbf{a}^1\|^2 - 1}$$

Case II: $a^1_j = 0$

$$0 = -w_{i,j}^{1:2}\|\mathbf{a}^1\|^2 \Rightarrow w_{i,j}^{1:2} = 0$$

Summary

$${}_i\mathbf{w}^{1:2} = \frac{\zeta \mathbf{a}^1}{\zeta + \|\mathbf{a}^1\|^2 - 1}$$

نظریه‌ی تشدید وفقی

قانون یادگیری L1-L2: یادگیری سریع

FAST LEARNING

فرض می‌کنیم خروجی‌های لایه‌ی ۱ و ۲ تا رسیدن وزن‌ها به حالت ماندگار، ثابت باقی می‌مانند ⇐ **یادگیری سریع**

$$\frac{dw_{i,j}^{1:2}(t)}{dt} = a_i^2(t)\left[(1-w_{i,j}^{1:2}(t))\zeta a_j^1(t) - w_{i,j}^{1:2}(t)\sum_{k\neq j} a_k^1(t)\right]$$

For *fast learning* we assume that the outputs of Layer 1 and Layer 2 remain constant until the weights reach steady state.

Assume that $a^2{}_i(t) = 1$, and solve for the steady state weight:

با فرض فعال بودن نرون i در لایه‌ی ۲:

مشتق مساوی صفر در حالت ماندگار $0 = \left[(1-w_{i,j}^{1:2})\zeta a_j^1 - w_{i,j}^{1:2}\sum_{k\neq j} a_k^1\right]$

Case I: $a^1{}_j = 1$

$$0 = (1-w_{i,j}^{1:2})\zeta - w_{i,j}^{1:2}(\|\mathbf{a}^1\|^2 - 1) = -(\zeta + \|\mathbf{a}^1\|^2 - 1)w_{i,j}^{1:2} + \zeta \Rightarrow w_{i,j}^{1:2} = \frac{\zeta}{\zeta + \|\mathbf{a}^1\|^2 - 1}$$

Case II: $a^1{}_j = 0$

$$0 = -w_{i,j}^{1:2}\|\mathbf{a}^1\|^2 \Rightarrow w_{i,j}^{1:2} = 0$$

Summary

$$_i\mathbf{w}^{1:2} = \frac{\zeta \mathbf{a}^1}{\zeta + \|\mathbf{a}^1\|^2 - 1}$$

$\zeta > 1$ می‌باشد تا اطمینان حاصل شود که مخرج کسر هرگز صفر نمی‌شود.

⇐ الگوهای پروتوتایپ نرمال* می‌شوند (حل مشکل زیرمجموعه/ابرمجموعه)

* در اینجا منظور از نرمال بودن، داشتن اندازه‌ی ۱ نیست، بلکه منظور این است که سطرهایی از $\mathbf{W}^{1:2}$ که اعضای غیر صفر بیشتری دارند، اندازه‌ی کوچکتری نسبت به سایرین دارند.

Learning Law: L2-L1

Outstar

$$\frac{d[\mathbf{w}_j^{2:1}(t)]}{dt} = a_j^2(t)[-\mathbf{w}_j^{2:1}(t) + \mathbf{a}^1(t)]$$

Fast Learning

Assume that $a^2_j(t) = 1$, and solve for the steady state weight:

$$\mathbf{0} = -\mathbf{w}_j^{2:1} + \mathbf{a}^1 \quad \text{or} \quad \mathbf{w}_j^{2:1} = \mathbf{a}^1$$

Column j of $\mathbf{W}^{2:1}$ converges to the output of Layer 1, which is a combination of the input pattern and the previous prototype pattern. The prototype pattern is modified to incorporate the current input pattern.

نظریه‌ی تشدید وفقی

قانون یادگیری L2-L1

LEARNING LAW: L2-L1

در معماری ART1، اتصالات L2-L1 با استفاده از قاعده‌ی یادگیری Outstar آموزش داده می‌شوند.

Outstar

$$\frac{d[\mathbf{w}_j^{2:1}(t)]}{dt} = a_j^2(t)[-\mathbf{w}_j^{2:1}(t) + \mathbf{a}^1(t)]$$

در صورتی که نرون j در لایه‌ی ۲ فعال باشد (برنده‌ی رقابت)، آن‌گاه ستون jام $\mathbf{W}^{2:1}$ به‌سمت الگوی $\mathbf{a}^1$ حرکت می‌کند.

Fast Learning

با فرض سناریوی یادگیری سریع (ثابت ماندن خروجی لایه‌های ۱ و ۲ تا رسیدن وزن‌ها به حالت ماندگار):

Assume that $a_j^2(t) = 1$, and solve for the steady state weight:

فرض می‌کنیم که نرون j در لایه‌ی ۲ فعال باشد، یعنی $a_j^2 = 1$ و معادله را برای وزن حالت ماندگار حل می‌کنیم:

مشتق مساوی صفر در حالت ماندگار $\quad \mathbf{0} = -\mathbf{w}_j^{2:1} + \mathbf{a}^1 \quad \text{or} \quad \mathbf{w}_j^{2:1} = \mathbf{a}^1$

Column j of $\mathbf{W}^{2:1}$ converges to the output of Layer 1, which is a combination of the input pattern and the previous prototype pattern. The prototype pattern is modified to incorporate the current input pattern.

پس ستون jام $\mathbf{W}^{2:1}$ به‌سمت خروجی لایه‌ی ۱، $\mathbf{a}^1$، همگرا می‌شود (که ترکیبی از الگوی ورودی و پروتوتایپ قبلی است).
پس الگوی پروتوتایپ تغییر داده می‌شود تا الگوی ورودی جاری را دربرگیرد.

نظریه‌ی تشدید وفقی

قانون یادگیری L2-L1

LEARNING LAW: L2-L1

در معماری ART1، اتصالات L2-L1 با استفاده از قاعده‌ی یادگیری Outstar آموزش داده می‌شوند.

هدف از اتصالات L2-L1 به‌یادآوردن الگوهای پروتوتایپ مناسب (انتظارها) است.
پس می‌توان این الگوهای پروتوتایپ را با الگوهای ورودی در لایه‌ی ۱ مقایسه و ترکیب نمود.

زمانی که انتظار و الگوی ورودی از انطباق کافی برخوردار نیستند،
یک سیگنال reset به لایه‌ی ۲ ارسال می‌شود.
⇐ یک الگوی پروتوتایپ جدید قابلیت انتخاب شدن پیدا می‌کند.

❋ $\mathbf{W}^{1:2}$ و $\mathbf{W}^{2:1}$ همزمان به‌روزرسانی می‌شوند:

وقتی نرون j–ام لایه‌ی ۲ فعال است و انطباق کافی بین انتظار و الگوی ورودی وجود دارد (شرایط تشدید)،
آن‌گاه سطر j–ام از $\mathbf{W}^{1:2}$ و ستون j–ام از $\mathbf{W}^{2:1}$ وفق داده می‌شوند (وفق‌دهی):

ستون j–ام از $\mathbf{W}^{2:1}$ مساوی $\mathbf{a}^1$ قرار داده می‌شود و به‌طور همزمان
سطر j–ام از $\mathbf{W}^{1:2}$ برابر با نسخه‌ی نرمال‌شده‌ی $\mathbf{a}^1$ قرار داده می‌شود.

شبکه‌های عصبی مصنوعی

نظریه‌ی تشدید وفقی

۶

خلاصه‌ی الگوریتم ART1

نظریه‌ی تشدید وفقی

خلاصه‌ی الگوریتم ART1 : مقداردهی آغازین

ART1 ALGORITHM SUMMARY: INITIALIZATION

مقداردهی آغازین ماتریس‌های وزن:

اعضای ماتریس $\mathbf{W}^{2:1}$ همگی برابر با 1 قرار داده می‌شوند.

$\Downarrow$

اولین باری که یک نرون جدید از لایه‌ی ۲ در رقابت برنده می‌شود، **تشدید** رخ می‌دهد،

زیرا: $\mathbf{a}^1 = \mathbf{p} \cap \mathbf{w}_j^{2:1} = \mathbf{p}$ و در نتیجه

$$\|\mathbf{a}^1\|^2 / \|\mathbf{p}\|^2 = 1 > \rho$$

$\Downarrow$

هر ستون آموزش ندیده در $\mathbf{W}^{2:1}$ حالت خام دارد

و قابلیت وفق‌یابی با هر یک از الگوهای ورودی را داراست.

از آنجا که سطرهای $\mathbf{W}^{1:2}$ باید نسخه‌ی نرمال‌شده‌ی ستون‌های $\mathbf{W}^{2:1}$ باشد،

برای عناصر آن قرار می‌دهیم:

$$\frac{\zeta}{\zeta + S^1 - 1}$$

ART1 Algorithm Summary

0) All elements of the initial $\mathbf{W}^{2:1}$ matrix are set to 1. All elements of the initial $\mathbf{W}^{1:2}$ matrix are set to $\zeta/(\zeta+S^1-1)$.

1) Input pattern is presented. Since Layer 2 is not active,

$$\mathbf{a}^1 = \mathbf{p}$$

2) The input to Layer 2 is computed, and the neuron with the largest input is activated.

$$a_i^2 = \begin{cases} 1, & \text{if}({}_i\mathbf{w}^{1:2})^T\mathbf{a}^1 = max[({}_k\mathbf{w}^{1:2})^T\mathbf{a}^1]) \\ 0, & \text{otherwise} \end{cases}$$

In case of a tie, the neuron with the smallest index is the winner.

3) The L2-L1 expectation is computed.

$$\mathbf{W}^{2:1}\mathbf{a}^2 = \mathbf{w}_j^{2:1}$$

نظریه‌ی تشدید وفقی

خلاصه‌ی الگوریتم ART1 (۱ از ۲)

ART1 ALGORITHM SUMMARY

0) All elements of the initial $\mathbf{W}^{2:1}$ matrix are set to 1. All elements of the initial $\mathbf{W}^{1:2}$ matrix are set to $\zeta/(\zeta+S^1-1)$.

1) Input pattern is presented. Since Layer 2 is not active,

نمایش الگوی ورودی : $\mathbf{a}^1 = \mathbf{p}$

2) The input to Layer 2 is computed, and the neuron with the largest input is activated.

محاسبه‌ی ورودی لایه‌ی ۲ : فعال شدن نرون دارای بزرگ‌ترین ورودی

$$a_i^2 = \begin{cases} 1, & \text{if}\left({}_i\mathbf{w}^{1:2}\right)^T\mathbf{a}^1 = max\left[\left({}_k\mathbf{w}^{1:2}\right)^T\mathbf{a}^1\right] \\ 0, & \text{otherwise} \end{cases}$$

In case of a tie, the neuron with the smallest index is the winner.

در صورتی که ورودی‌ها مساوی بودند، نرون برنده آنی است که کوچک‌ترین اندیس را دارد.

3) The L2-L1 expectation is computed.

محاسبه‌ی انتظار L2-L1 : $\mathbf{W}^{2:1}\mathbf{a}^2 = \mathbf{w}_j^{2:1}$

Summary Continued

4) Layer 1 output is adjusted to include the L2-L1 expectation.

$$\mathbf{a}^1 = \mathbf{p} \cap \mathbf{w}_j^{2:1}$$

5) The orienting subsystem determines match between the expectation and the input pattern.

$$a^0 = \begin{cases} 1, & \text{if}[\|\mathbf{a}^1\|^2 / \|\mathbf{p}\|^2 < \rho] \\ 0, & \text{otherwise} \end{cases}$$

6) If $a^0 = 1$, then set $a^2_j = 0$, inhibit it until resonance, and return to Step 1. If $a^0 = 0$, then continue with Step 7.

7) Resonance has occured. Update row j of $\mathbf{W}^{1:2}$.

$${}_j\mathbf{w}^{1:2} = \frac{\zeta \mathbf{a}^1}{\zeta + \|\mathbf{a}^1\|^2 - 1}$$

8) Update column j of $\mathbf{W}^{2:1}$.

$$\mathbf{w}_j^{2:1} = \mathbf{a}^1$$

9) Remove input, restore inhibited neurons, and return to Step 1.

نظریه‌ی تشدید وفقی

خلاصه‌ی الگوریتم ART1 (۲ از ۲)

ART1 ALGORITHM SUMMARY

4) Layer 1 output is adjusted to include the L2-L1 expectation.

تنظیم خروجی لایه‌ی ۱ برای وارد کردن انتظار : $\mathbf{a}^1 = \mathbf{p} \cap \mathbf{w}_j^{2:1}$

5) The orienting subsystem determines match between the expectation and the input pattern.

تعیین میزان انطباق انتظار با الگوی ورودی توسط زیرسیستم گرادهنده : $a^0 = \begin{cases} 1, & \text{if}[\|\mathbf{a}^1\|^2 / \|\mathbf{p}\|^2 < \rho] \\ 0, & \text{otherwise} \end{cases}$

6) If $a^0 = 1$, then set $a^2_j = 0$, inhibit it until resonance, and return to Step 1. If $a^0 = 0$, then continue with Step 7.

انطباق کافی

7) Resonance has occured. Update row j of $\mathbf{W}^{1:2}$.

وقوع تشدید: ⇐ به‌هنگام‌سازی وزن‌ها

$$_j\mathbf{w}^{1:2} = \frac{\zeta \mathbf{a}^1}{\zeta + \|\mathbf{a}^1\|^2 - 1}$$

8) Update column j of $\mathbf{W}^{2:1}$.

$$\mathbf{w}_j^{2:1} = \mathbf{a}^1$$

اعمال الگوهای ورودی به شبکه تا رسیدن به وزن‌های ماندگار

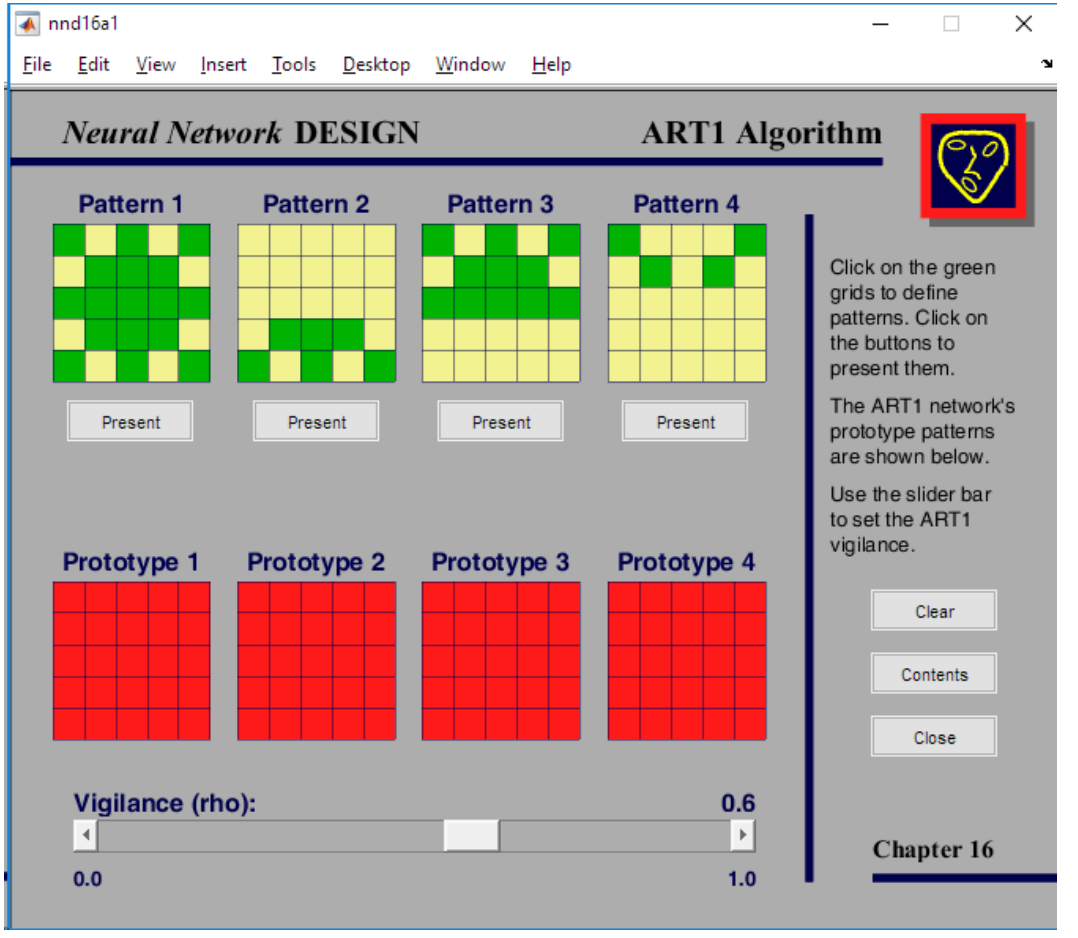
9) Remove input, restore inhibited neurons, and return to Step 1.

نظریه‌ی تشدید وفقی

همگرایی الگوریتم ART1

ART1 ALGORITHM CONVERGENCE

کارپنتر و گراسبرگ ثابت کردند که
الگوریتم ART1 به‌ازای هر مجموعه از الگوهای ورودی **خوشه‌های پایدار** تشکیل می‌دهد.


nnd16a1
File Edit View Insert Tools Desktop Window Help
Neural Network DESIGN
ART1 Algorithm
Pattern 1
Pattern 2
Pattern 3
Pattern 4
Present
Present
Present
Present
Click on the green grids to define patterns. Click on the buttons to present them.
The ART1 network's prototype patterns are shown below.
Use the slider bar to set the ART1 vigilance.
Prototype 1
Prototype 2
Prototype 3
Prototype 4
Clear
Contents
Close
Vigilance (rho):
0.6
0.0
1.0
Chapter 16

```
>> nnd16a1
```

شبکه‌های عصبی مصنوعی

نظریه‌ی تشدید وفقی

۷

دیگر معماری‌های ART

دیگر معماری‌های ART

OTHER ART ARCHITECTURES

یک محدودیت ART1: فقط بر روی الگوهای ورودی دودویی قابل تعریف است.

ART1 تنها یکی از مثال‌های نظریه‌ی تشدید وفقی است.
کارپنتر، گراسبرگ و افراد دیگر، شکل‌های گوناگون دیگری از شبکه‌ی ART را پیشنهاد داده‌اند.

شبکه‌ی ART2
ART2 Network

شبکه‌ی ART3
ART3 Network

شبکه‌ی ARTMAP
ARTMAP Network

شبکه‌ی ARTMAP فازی
Fuzzy ARTMAP Network

دیگر معماری‌های ART

شبکه‌ی ART2

OTHER ART ARCHITECTURES

شبکه‌ی ART2
ART2 Network

[کارپنتر و گراسبرگ]

برای استفاده در مورد الگوهای دودویی و آنالوگ

اصول ساختاری ART2 بسیار مشابه ART1 است.

تفاوت:
لایه‌ی ۱ شبکه‌ی ART2 دارای چندین زیرلایه است.

(زیرا بردارهای آنالوگ برخلاف بردارهای دودویی تا هر اندازه‌ای می‌توانند به هم نزدیک باشند؛
وجود این زیرلایه‌ها ضروری است زیرا این زیرلایه‌ها ترکیبی از اعمال نرمال‌سازی و حذف نویز را
علاوه بر مقایسه‌ی بردار ورودی و انتظار انجام می‌دهند.)

دیگر معماری‌های ART

شبکه‌ی ART3

OTHER ART ARCHITECTURES

شبکه‌ی ART3
ART3 Network

[کارپنتر و گراسبرگ]

دارای مکانیسم بازنشانی (reset) پیچیده‌تر [نزدیک‌تر به مدل‌های بیولوژیکی]

دیگر معماری‌های ART

شبکه‌ی ARTMAP

OTHER ART ARCHITECTURES

شبکه‌ی ARTMAP
ARTMAP Network

[کارپنتر، گراسبرگ و رینولدز]

یک شبکه‌ی بانظارت (برخلاف انواع دیگر ART)

شامل دو ماژول ART که توسط یک حافظه‌ی انجمنی inter-ART به هم متصل شده‌اند.

یک ماژول ART: برای دریافت بردار ورودی (input)
یک ماژول ART دیگر: برای دریافت بردار خروجی مطلوب (target)

شبکه یاد می‌گیرد با ارائه‌ی یک بردار ورودی به آن، بردار خروجی صحیح را پیش‌بینی کند.

دیگر معماری‌های ART

شبکه‌ی ARTMAP فازی

OTHER ART ARCHITECTURES

شبکه‌ی ARTMAP فازی
Fuzzy ARTMAP Network

[کارپنتر، گراسبرگ، مارکوزن، رینولدز و روزن]

استفاده از منطق فازی در ساختار شبکه‌ی ART

افزایش کارآیی به شکل قابل توجه
به‌خصوص در مورد الگوهای ورودی دارای نویز

دیگر معماری‌های ART

مشترکات

OTHER ART ARCHITECTURES

همه‌ی معماری‌های مختلف ART ماژول‌های کلیدی ART1 را دارند، از جمله:

- **قاعده‌ی Instar** در اتصالات L1-L2 برای بازشناسی الگوها
- **قاعده‌ی Outstar** در اتصالات L2-L1 برای به‌یادآوری الگوها
- **لایه‌ی ۲** با هدف ارتقای کنتراست (رقابت)
- **لایه‌ی ۱** با هدف مقایسه‌ی الگوهای «ورودی» و «انتظار»
- **زیرسیستم گرادهنده** برای بازنشانی در هنگام عدم تطابق الگوها

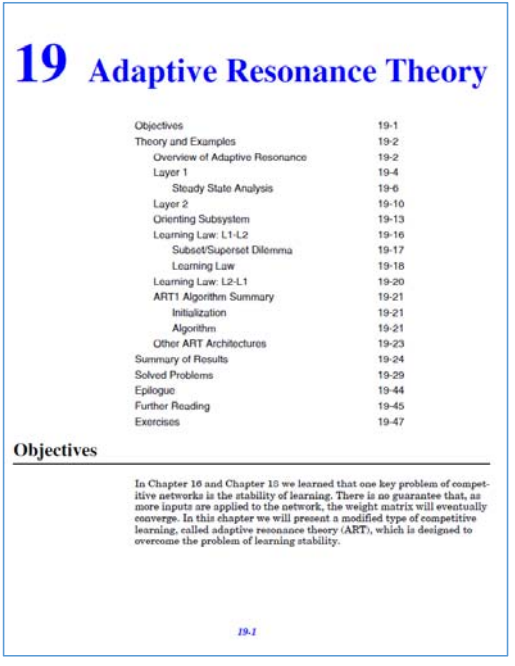
شبکه‌های عصبی مصنوعی

نظریه‌ی تشدید وفقی

۸

منابع

منبع اصلی



19 Adaptive Resonance Theory

Objectives 19-1
Theory and Examples 19-2
Overview of Adaptive Resonance 19-2
Layer 1 19-4
Steady State Analysis 19-6
Layer 2 19-10
Orienting Subsystem 19-13
Learning Law: L1-L2 19-16
Subset/Superset Dilemma 19-17
Learning Law 19-18
Learning Law: L2-L1 19-20
ART1 Algorithm Summary 19-21
Initialization 19-21
Algorithm 19-21
Other ART Architectures 19-23
Summary of Results 19-24
Solved Problems 19-29
Epilogue 19-44
Further Reading 19-45
Exercises 19-47

Objectives

In Chapter 16 and Chapter 18 we learned that one key problem of competitive networks is the stability of learning. There is no guarantee that, as more inputs are applied to the network, the weight matrix will eventually converge. In this chapter we will present a modified type of competitive learning, called adaptive resonance theory (ART), which is designed to overcome the problem of learning stability.

19-1

Martin T. Hagan, Howard B. Demuth, Mark H. Beale, Orlando De Jesus,
Neural Network Design,
2nd Edition, Martin Hagan, 2014.
Chapter 19

Online version can be downloaded from: http://hagan.okstate.edu/nnd.html