# شبکه‌های عصبی مصنوعی

## درس ۱۸

# شبکه‌ی گراسبرگ

## Grossberg Network

کاظم فولادی قلعه
دانشکده مهندسی، پردیس فارابی
دانشگاه تهران

http://courses.fouladi.ir/nn

# Grossberg Network

# شبکه‌ی گراسبرگ

GROSSBERG NETWORK

**شبکه‌ی گراسبرگ:** در ادامه‌ی مباحث مربوط به الگوریتم‌های یادگیری رقابتی و انجمنی

**شبکه‌ی گراسبرگ:** نوعی شبکه‌ی رقابتی خود–سازمان‌ده پیوسته–زمان؛
با الهام از سیستم بینایی انسان؛
پایه و اساس شبکه‌های نظریه‌ی تشدید وقفی (ART)

گراسبرگ، از پژوهشگران پرکار دوره‌ی افت پژوهش بر روی شبکه‌های عصبی (اواخر 1960 و اوایل 1970) بود.
گراسبرگ، از ریاضیات غیرخطی برای مدل‌سازی کارکردهای خاص مغزی استفاده می‌کرد.

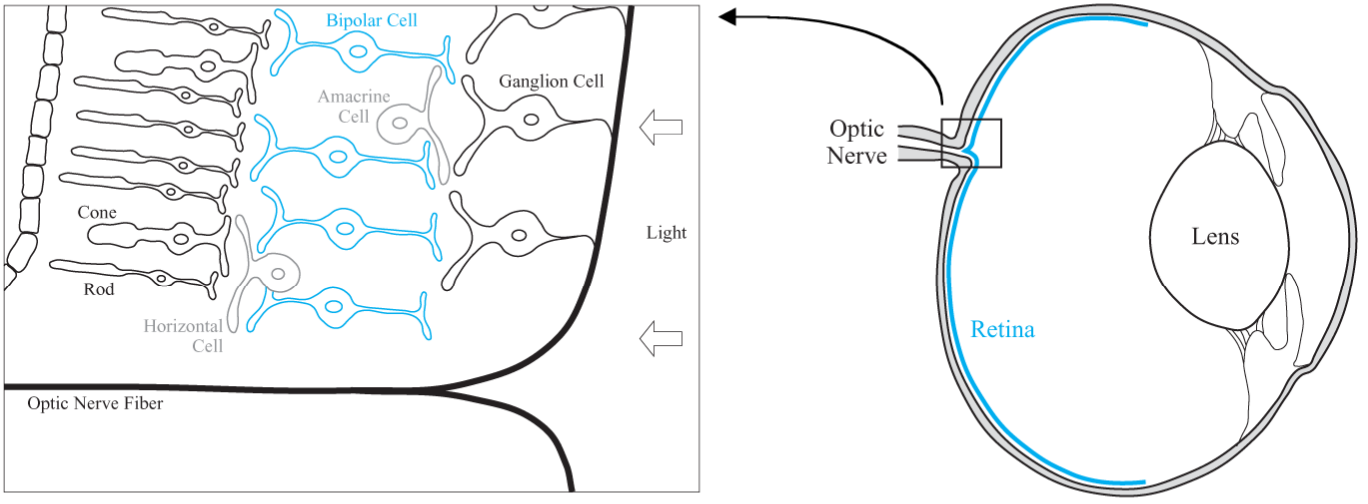
**نکته‌ی مهم:**
حوزه‌ی شبکه‌های عصبی، مخلوطی از زمینه‌های زیست‌شناسی، ریاضی و روان‌شناسی است و شناخت دقیق مفاهیم شبکه‌های عصبی نیازمند کنکاش در چنین زمینه‌هایی است.

شبکه‌ی گراسبرگ

# ۱

# انگیزش بیولوژیکی: بینایی

# Biological Motivation: Vision



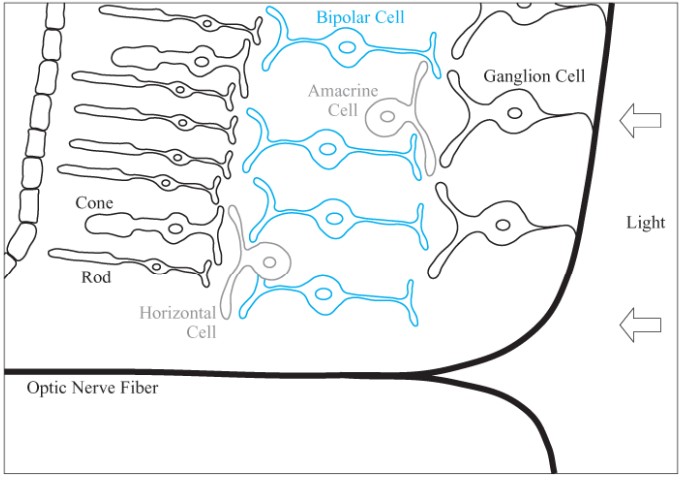

Eyeball and Retina

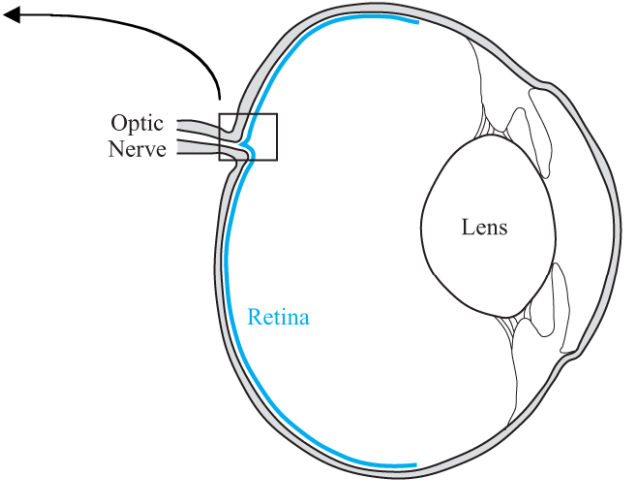
# سیستم بینایی انسان

انگیزش بیولوژیکی شبکه‌های گراسبرگ

## BIOLOGICAL MOTIVATION: VISION

نور از قرنیه و عدسی عبور کرده و توسط عدسی بر روی شبکیه متمرکز می‌شود.




پس از رسیدن نور به شبکیه،
فرآیند گسترده‌ی ترجمه‌ی اطلاعات نوری به یک تصویر قابل فهم برای انسان آغاز می‌شود.

Eyeball and Retina

# سیستم بینایی انسان

## شبکیه

## RETINA

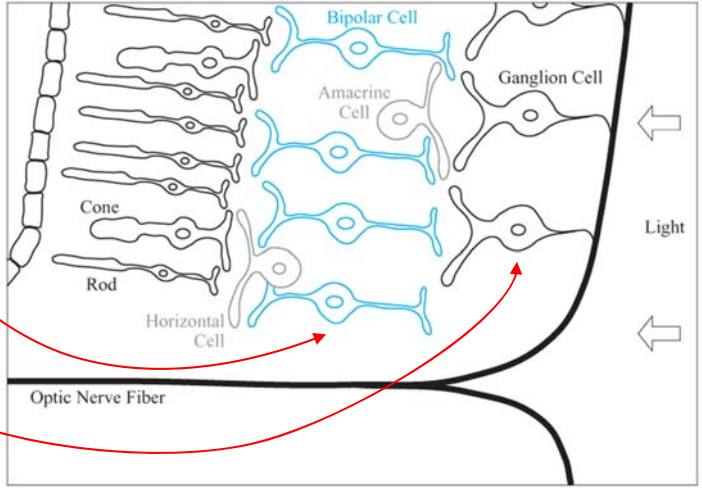
شبکیه، در واقع قسمتی از مغز می‌باشد که در طی مراحل رشد جنین از مغز جدا می‌شود، اما اتصال بین شبکیه و مغز از طریق عصب‌های بینایی برقرار است.

شبکیه دیواره‌ی داخلی پشت چشم را می‌پوشاند و از سه لایه نرون تشکیل شده است:

**لایه‌ی خارجی**
*Outer Layer*

**لایه‌ی میانی**
*Middle Layer*

**لایه‌ی نهایی**
*Final Layer*



# Layers of Retina

The retina is a part of the brain that covers the back inner wall of the eye and consists of three layers of neurons:

Outer Layer:

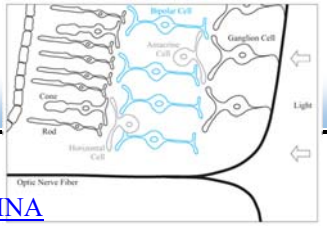
- Photoreceptors - convert light into electrical signals
  - Rods - allow us to see in dim light
  - Cones - fine detail and color

Middle Layer

- Bipolar Cells - link photoreceptors to third layer
- Horizontal Cells - link receptors with bipolar cells
- Amacrine Cells - link bipolar cells with ganglion cells

Final Layer

- Ganglion Cells - link retina to brain through optic nerve




# سیستم بینایی انسان

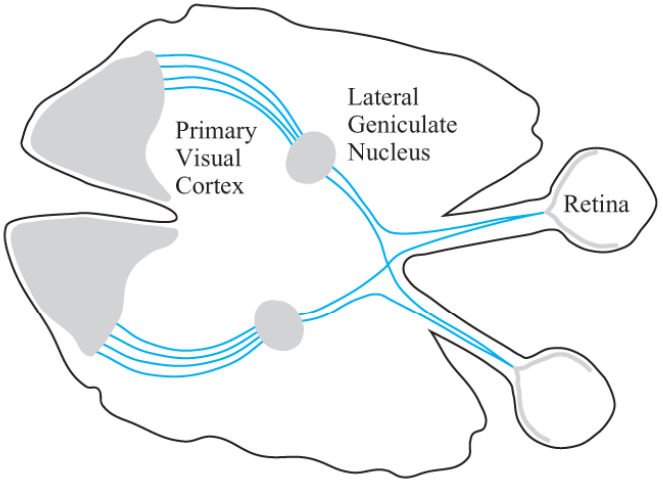
## لایه‌های شبکیه

## LAYERS OF RETINA

| | |
|---|---|
| **لایه‌ی خارجی** *Outer Layer* | شامل سلول‌های گیرنده‌ی نور: تبدیل نور به سیگنال‌های الکتریکی |
| **گیرنده‌های نور** *Photoreceptors* | **سلول‌های میله‌ای** (*Rods*): توانایی دیدن نور تاریک (dim light) |
| | **سلول‌های مخروطی** (*Cons*): توانایی دیدن جزئیات ریز و رنگ‌ها |
| **لایه‌ی میانی** *Middle Layer* | به‌دلایل ناشناخته نور باید از دو لایه‌ی دیگر از شبکیه عبور کند تا سلول‌های میله‌ای و مخروطی تحریک شوند: لایه‌ی میانی و لایه‌ی نهایی |
| **سلول‌های دوقطبی** (*Bipolar Cells*) | اتصال‌دهنده‌ی گیرنده‌های نور به لایه‌ی نهایی |
| **سلول‌های افقی** (*Horizontal Cells*) | اتصال‌دهنده‌ی بین گیرنده‌ها و سلول‌های دوقطبی |
| **سلول‌های آماکراین** (*Amacrine Cells*) | اتصال‌دهنده‌ی سلول‌های دوقطبی با سلول‌های لنفاوی لایه‌ی نهایی |
| **لایه‌ی نهایی** *Final Layer* | متشکل از سلول‌های لنفاوی: آکسون سلول‌های لنفاوی از سطح شبکیه عبور کرده و در کنار هم جمع شده‌اند و عصب‌های بینایی را شکل داده‌اند. |
| **سلول‌های لنفاوی** (*Ganglion Cells*) | اتصال‌دهنده‌ی شبکیه به مغز از طریق عصب اپتیکی |

هر چشم از حدود ۱۲۵ میلیون سلول گیرنده‌ی نور تشکیل می‌شود، اما فقط ۱ میلیون سلول لنفاوی دارد
⇐ فرآیندهای قابل توجهی در شبکیه در راستای **کاهش داده‌ها** انجام می‌شود.

# Visual Pathway

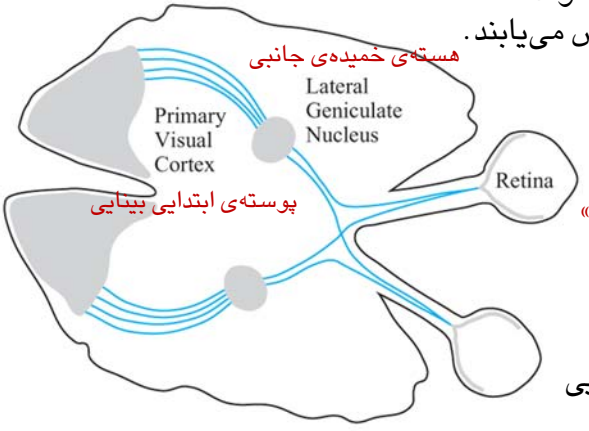


# سیستم بینایی انسان

## مسیر بینایی

## VISUAL PATHWAY

آکسون‌های سلول‌های لنفاوی در قالب عصب‌های بینایی جمع شده و به قسمتی از مغز تحت عنوان «هسته‌ی خمیده‌ی جانبی» متصل می‌شوند.
از این نقطه به بعد فیبرها بر روی پوسته‌ی ابتدایی بینایی گسترش می‌یابند.

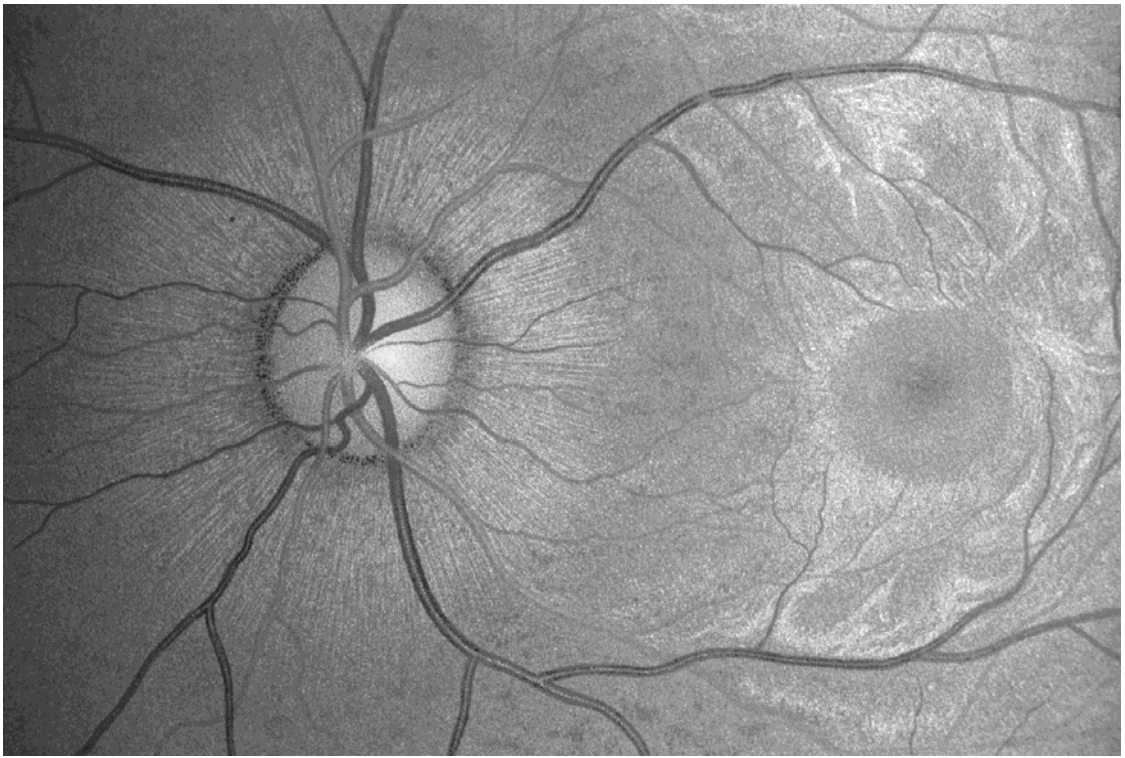



آکسون‌های سلول‌های لنفاوی با «سلول‌های خمیده‌ی جانبی» تشکیل سیناپس می‌دهند.
همچنین «سلول‌های خمیده‌ی جانبی» با «پوسته‌ی ابتدایی بینایی» تشکیل سیناپس می‌دهند.

«پوسته‌ی بینایی» قسمتی از مغز است که مختص به کارکرد بینایی بوده و شامل لایه‌های متعدد است.

- اتصالات بین مسیرهای بینایی کاملاً غیرتصادفی است؛
  نگاشت هر لایه به لایه‌ی بعدی کاملاً سازمان‌دهی‌شده است.
- آکسون‌های سلول‌های لنفاوی در قسمت خاصی از شبکیه به قسمت مشخصی از سلول‌های خمیده‌ی جانبی رفته و سپس به قسمت مشخصی از قشر بینایی متصل می‌شوند.
- به‌علاوه هر نیم‌کره اطلاعات ورودی را از هر دو چشم دریافت می‌کند
  (زیرا برخی فیبرها از روی هم می‌گذرند و برخی بدون تقاطع هستند).

# Photograph of the Retina

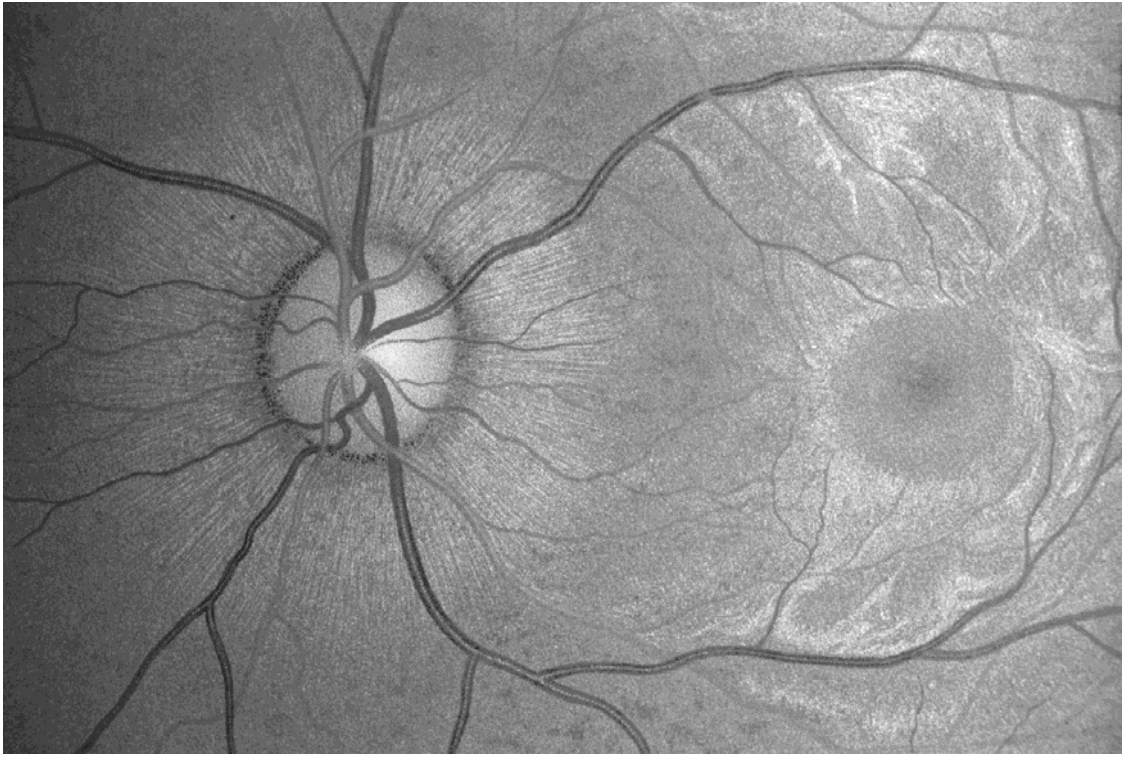


# سیستم بینایی انسان

عکس شبکیه

## PHOTOGRAPH OF THE RETINA

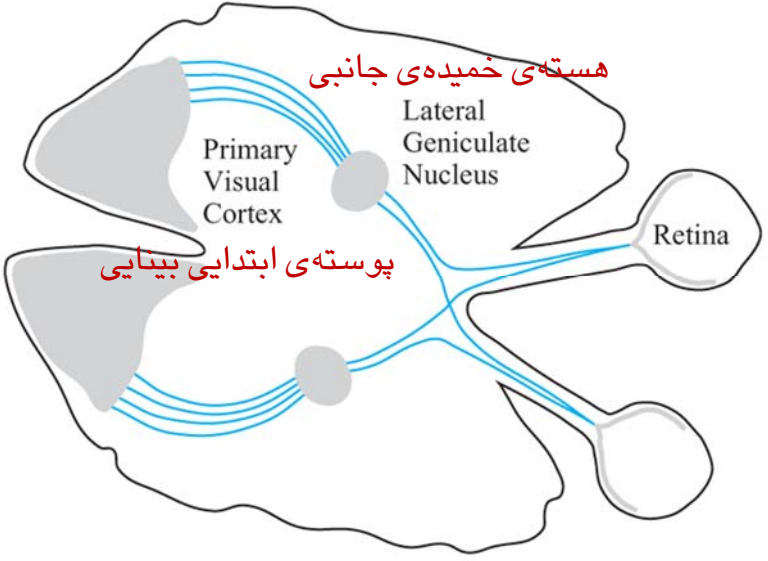
Blind Spot (Optic Disk) نقطه‌ی کور



# سیستم بینایی انسان

مسیر بینایی

## VISUAL PATHWAY




اکنون یک ایده‌ی کلی در مورد ساختار و مسیرهای بینایی دریافت کرده‌ایم. اما:

ـ کارکرد بینایی چگونه انجام می‌شود؟

ـ هدف سه لایه‌ی شبکیه چیست؟

ـ چه عملیاتی توسط سلول‌های خمیده‌ی جانبی انجام می‌شود؟

برای پاسخ به این پرسش‌ها، باید خطاهای دید را بررسی کنیم.

# سیستم بینایی انسان

نواقص درک شبکیه

## IMPERFECTIONS IN RETINAL UPTAKE

علت وجود خطای دید چیست؟
نقص‌های موجود در فرآیند درک شبکیه موجب خطای دید می‌شود.

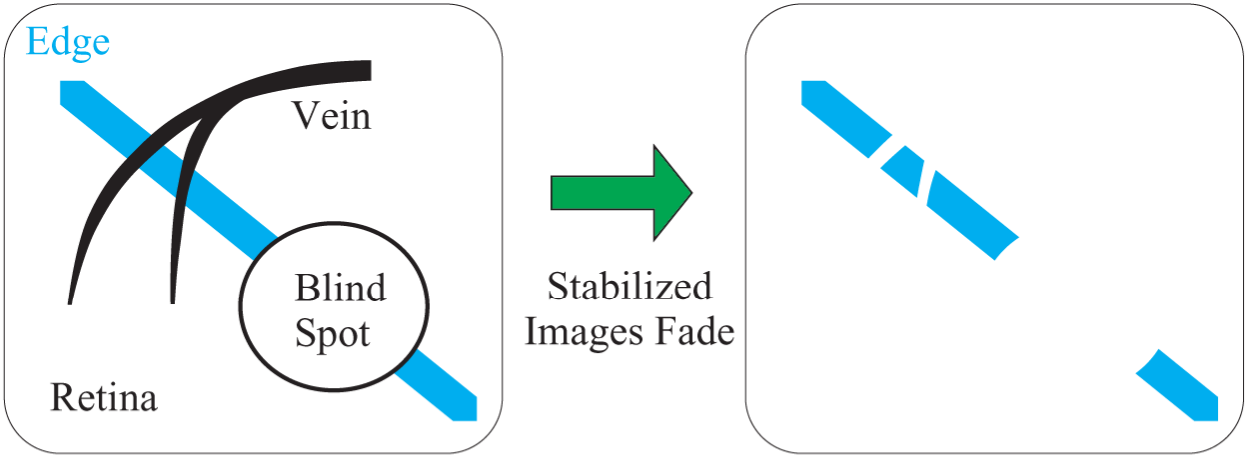
اگر بتوانیم یک مدل ریاضی را برای تولید خطاهای دید در سیستم بینایی توسعه دهیم،
آن‌گاه ممکن است بتوانیم نحوه‌ی عملکرد قسمتی از مغز را توجیه نماییم.

**نقص یک در سیستم ادراکی شبیکه:**
سلول‌های میله‌ای و مخروطی در نقطه‌ی کور وجود ندارند (علت شکل‌گیری نقطه‌ی کور)
ما به‌طور عادی متوجه وجود این نقطه‌ی کور در چشممان نیستیم (به‌دلیل فرآیندهایی که در مسیر بینایی رخ می‌دهد ).

آزمایش برای نقطه‌ی کور

به دایره‌ی سمت چپ خیره شوید تا زمانی که چشم چپ شما را به‌طور کامل پوشش دهد.
سپس با حرکت سر به این دایره نزدیک شوید
و سپس به آرامی از آن دور شوید.
متوجه یک نکته می‌شوید: دایره‌ی سمت راست از دید شما محو می‌شود.
(* نکته‌ی جالب: ما دخالتی در انتقال تصویر به نقطه‌ی کور نداریم بلکه مغز ماست که مرتکب این اشتباه می‌شود.)

# Imperfections in Retinal Uptake



# سیستم بینایی انسان

نواقص درک شبکیه

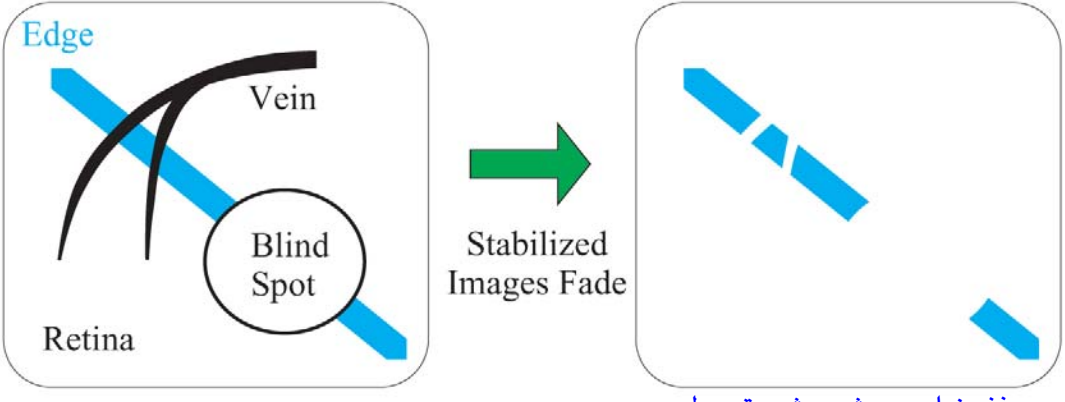
## IMPERFECTIONS IN RETINAL UPTAKE

### نقص دو در سیستم ادراکی شبیکه:

سرخرگ و سیاهرگ از مقابل گیرنده‌های نوری شبکیه عبور می‌کند.

به‌این ترتیب از دریافت نور توسط سلول‌های میله‌ای و مخروطی در ناحیه‌ی بینایی جلوگیری می‌شود؛
به‌علاوه، از آنجا که گیرنده‌های نور شبکیه پشت دو لایه از شبکیه قرار دارند، نور ابتدا باید از این دو لایه عبور کند.

از آنجا که این نقاط فاقد سلول‌های گیرنده‌ی نور هستند، ما نمی‌توانیم رگ‌های داخل چشم خودمان را ببینیم.
به‌علاوه، کره‌ی چشم به‌طور طبیعی دارای تکان‌های ثابتی است ⇐
حتی یک جسم ثابت نیز متناسب با حرکت کره‌ی چشم در ناحیه‌ی بینایی ما متحرک می‌باشد.
* از آنجا که رگ‌ها نسبت به کره‌ی چشم ثابت هستند، بنابراین از ناحیه‌ی دید ما محو می‌شوند.



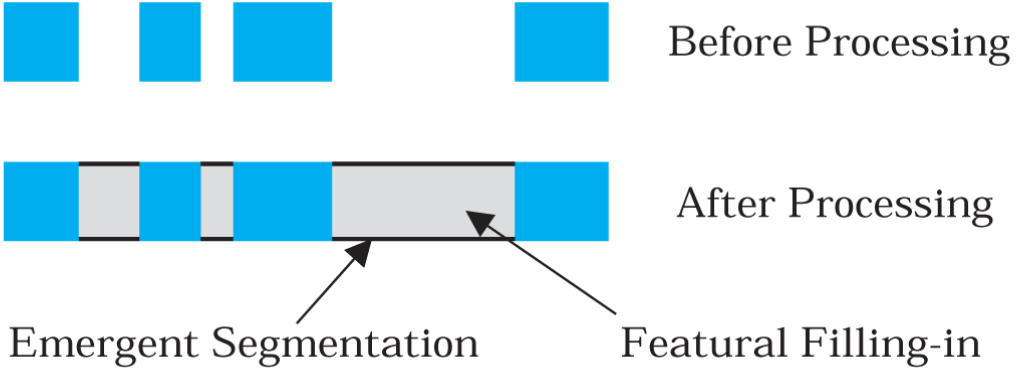

در این شکل لبه‌ی شبکیه نشان داده شده است.

حذف نواحی پوشیده شده توسط نقطه‌ی کور و رگ‌های چشمی از لبه‌ی شبکیه

# Compensatory Processing

**Emergent Segmentation**:
Complete missing boundaries.

**Featural Filling-In**:
Fill in color and brightness.



# سیستم بینایی انسان

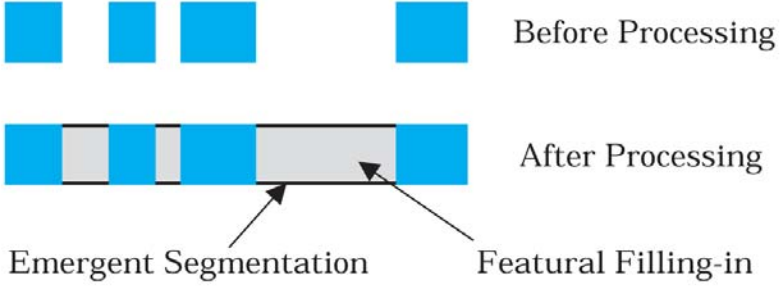
پردازش جبران‌ساز

## COMPENSATORY PROCESSING

از آنجا که ناحیه‌ی بینایی ما به‌دلیل وجود نقاط کور یاد شده دارای نقایص و کمبودهایی است، بنابراین سیستم عصبی ما باید این نقایص را با انجام یک سری عملیات خاص برطرف نماید و تصویر را کامل کند.
گراسبرگ در این ارتباط دو نوع اولیه از فرآیندهای جبران‌ساز را پیشنهاد کرده است:

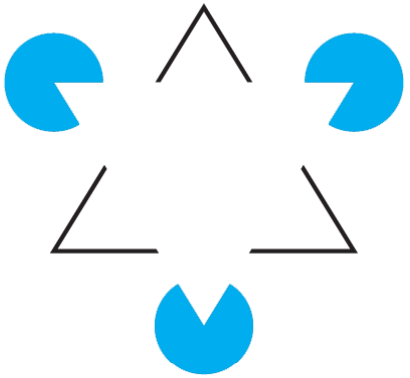
**Emergent Segmentation**:
Complete missing boundaries. قسمت‌های گم‌شده را تکمیل می‌کند.

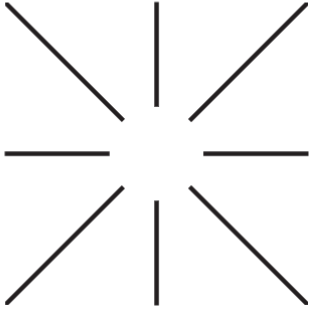
**Featural Filling-In**:
Fill in color and brightness. قسمت‌های تولید شده را رنگ‌آمیزی کرده و روشنایی آنها را تنظیم می‌کند.




از آنجا که فرآیند ترسیم تصویری که ما می‌بینیم نیاز به زمان دارد، بنابراین امکان تشخیص دقیق قسمت‌های گمشده‌ی تصویر وجود ندارد ⇐ **شکل‌گیری خطای دید**

# Visual Illusions





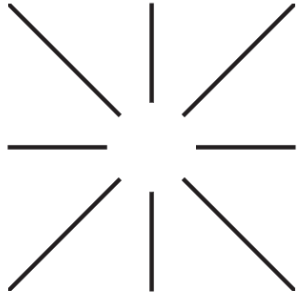
Illusions demostrate the compensatory processing of the visual system. Here we see a bright white triangle and a circle which do not actually exist in the figures.

# سیستم بینایی انسان

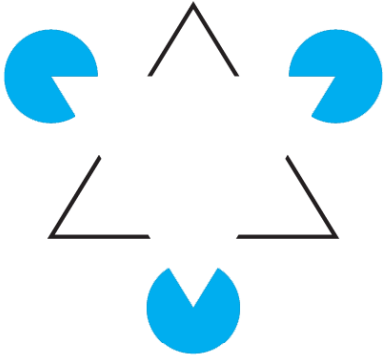
خطاهای دید

## VISUAL ILLUSIONS

در اثر دو فرآیند جبران‌سازی



یک دایره‌ی سفیدرنگ روی خطوط دیده می‌شود.

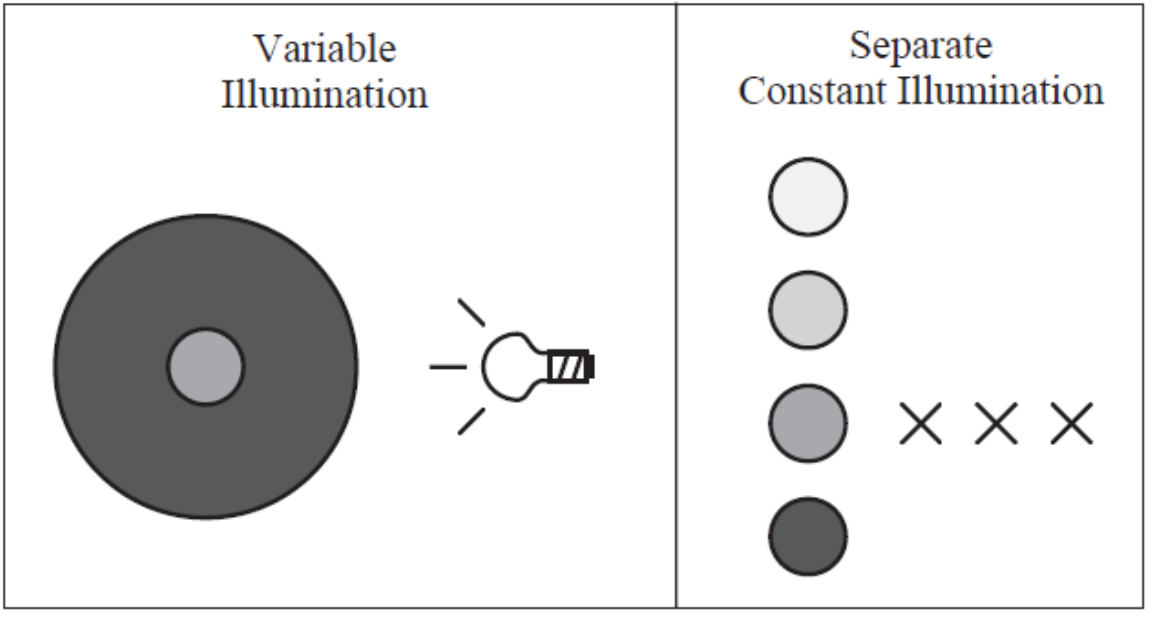


یک مثلث سفیدرنگ معکوس بر روی سه دایره‌ی آبی اطراف آن دیده می‌شود.

Illusions demostrate the compensatory processing of the visual system. Here we see a bright white triangle and a circle which do not actually exist in the figures.

علاوه بر دو فرآیند جبران‌ساز، دو پدیده‌ی دیگر هم به سیستم بینایی یاری می‌رسانند:
ثبات روشنایی + کنتراست روشنایی

# Vision Normalization



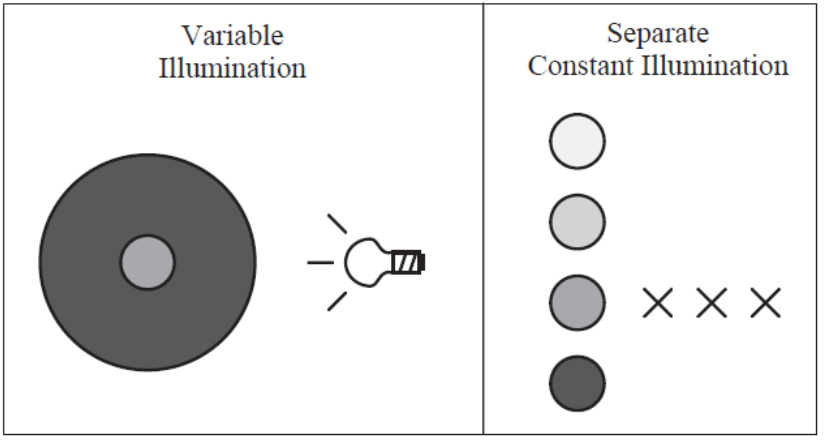

The vision systems normalize scenes so that we are only aware of relative differences in brightness, not absolute brightness.

# سیستم بینایی انسان

## نرمال‌سازی بینایی (تأثیر **ثبات روشنایی** در آزمایش)

## VISION NORMALIZATION (BRIGHTNESS CONSTANCY)

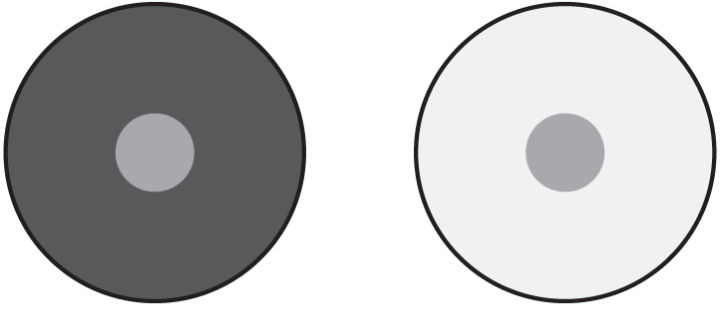
یک دیسک خاکستری که در داخل یک دیسک تیره‌تر قرار گرفته است، به فرد مورد آزمایش نشان داده می‌شود و از او خواسته می‌شود تا میزان روشنایی دیسک مرکزی را از طریق انتخاب یک دیسک هم‌رنگ در میان دیسک‌های موجود، تعیین نماید.




سپس میزان نوری که به تصویر می‌تابد افزایش داده می‌شود و دوباره از فرد خواسته می‌شود دیسک هم‌رنگ را انتخاب کند.
این فرآیند برای چند سطح نوری دیگر هم تکرار می‌شود.
در نتیجه‌ی این عملیات مشاهده می‌شود که فرد همیشه دیسک اول (دیسک درست) را انتخاب می‌کند.
این مورد در نتیجه‌ی فرآیند ثبات روشنایی است.

**سیستم بینایی صحنه‌ها را نرمال‌سازی می‌کند به‌گونه‌ای که ما فقط متوجه تغییرات روشنایی می‌شویم (نه مقدار مطلق آن).**

# Brightness Contrast



If you look at a point between the two circles, the small inner circle on the left will appear lighter than the small inner circle on the right, although they have the same brightness. It is relatively lighter than its surroundings.

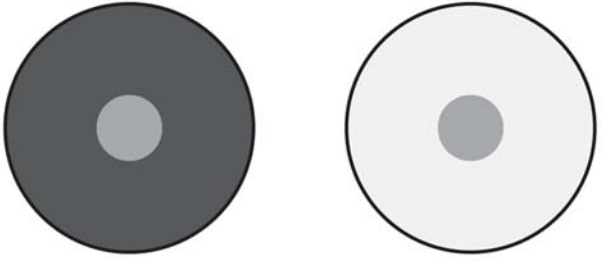
The visual system normalizes the scene. We see relative intensities.

# سیستم بینایی انسان

## نرمال‌سازی بینایی (کنتراست روشنایی)

## VISION NORMALIZATION (BRIGHTNESS CONTRAST)

هر چند هر دو دیسک میانی دارای روشنایی یکسان هستند، اما دیسک داخل دایره‌ی تیره، روشن‌تر به نظر می‌رسد.



دلیل این امر حساسیت سیستم بینایی ما به شدت نورهای مرتبط با هم می‌باشد.

**سیستم بینایی صحنه‌ها را نرمال‌سازی می‌کند به‌گونه‌ای که ما فقط متوجه شدت نورهای نسبی می‌شویم (نه مقدار مطلق آن).**

شبکه‌ی گراسبرگ

# ۲

# مدل غیرخطی پایه

## مدل غیرخطی پایه

BASIC NONLINEAR MODEL

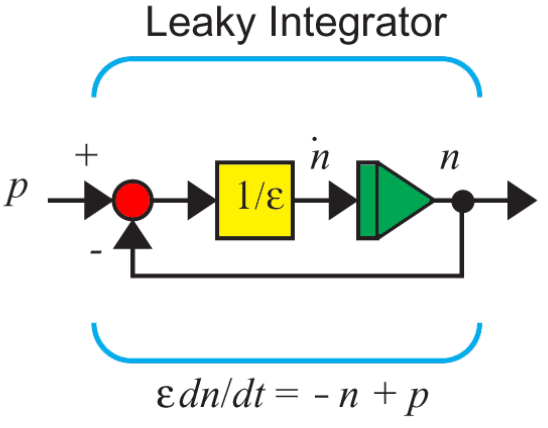
در ادامه:
ساختارهایی از شبکه‌های عصبی که با پدیده‌های فیزیکی تشریح شده
در مورد سیستم بینایی انسان سازگاری دارند،
را بررسی می‌کنیم.

برای آشنایی با شبکه‌ی گراسبرگ،
ابتدا باید برخی ساختارهای سازنده‌ی این شبکه را بررسی کنیم.

# Leaky Integrator

(Building block for basic nonlinear model.)

$$\varepsilon \frac{dn(t)}{dt} = -n(t) + p(t)$$



# مدل غیرخطی پایه

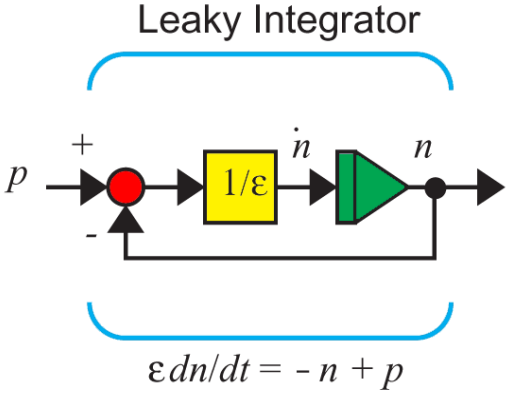
## انتگرال‌گیری نشتی‌دار

LEAKY INTEGRATOR

(Building block for basic nonlinear model.)

$$\varepsilon \frac{dn(t)}{dt} = -n(t) + p(t)$$

$\varepsilon$ ثابت زمانی سیستم است.

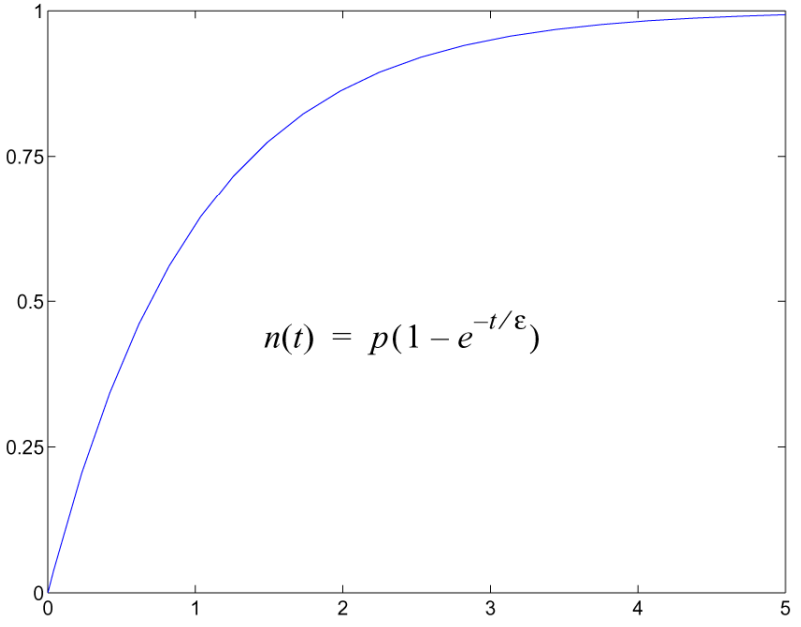



Prepared by Kazim Fouladi | Fall 2017 | 2nd Edition
کاظم فولادی

# Leaky Integrator Response

$$n(t) = e^{-t/\varepsilon} n(0) + \frac{1}{\varepsilon}\int_0^t e^{-(t-\tau)/\varepsilon} p(t-\tau)\,d\tau$$

For a constant input and zero initial conditions:



# مدل غیرخطی پایه

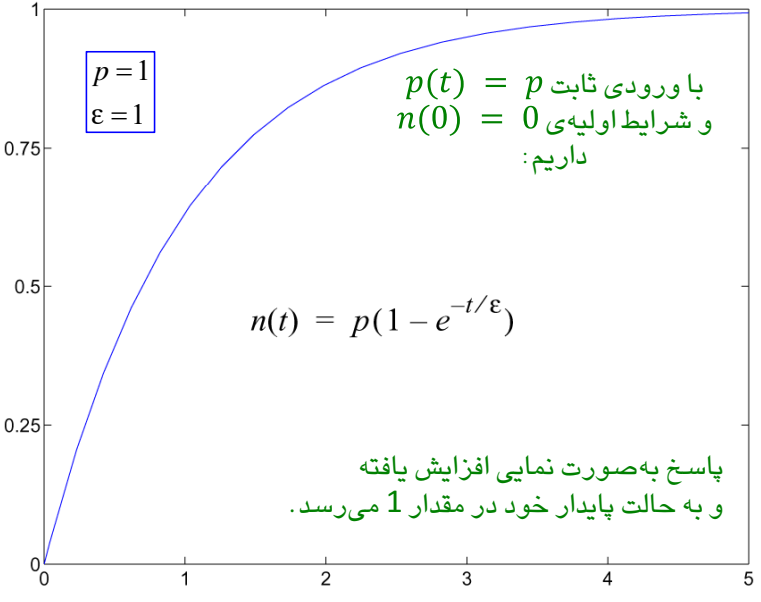
## پاسخ انتگرال‌گیری نشتی‌دار

## LEAKY INTEGRATOR RESPONSE

$$n(t) = e^{-t/\varepsilon} n(0) + \frac{1}{\varepsilon}\int_0^t e^{-(t-\tau)/\varepsilon} p(t-\tau)d\tau$$

پاسخ انتگرال‌گیر نشتی‌دار به ورودی دلخواه $p(t)$

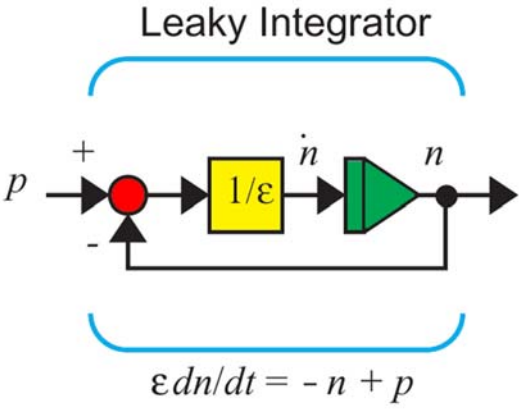
For a constant input and zero initial conditions:



# مدل غیرخطی پایه

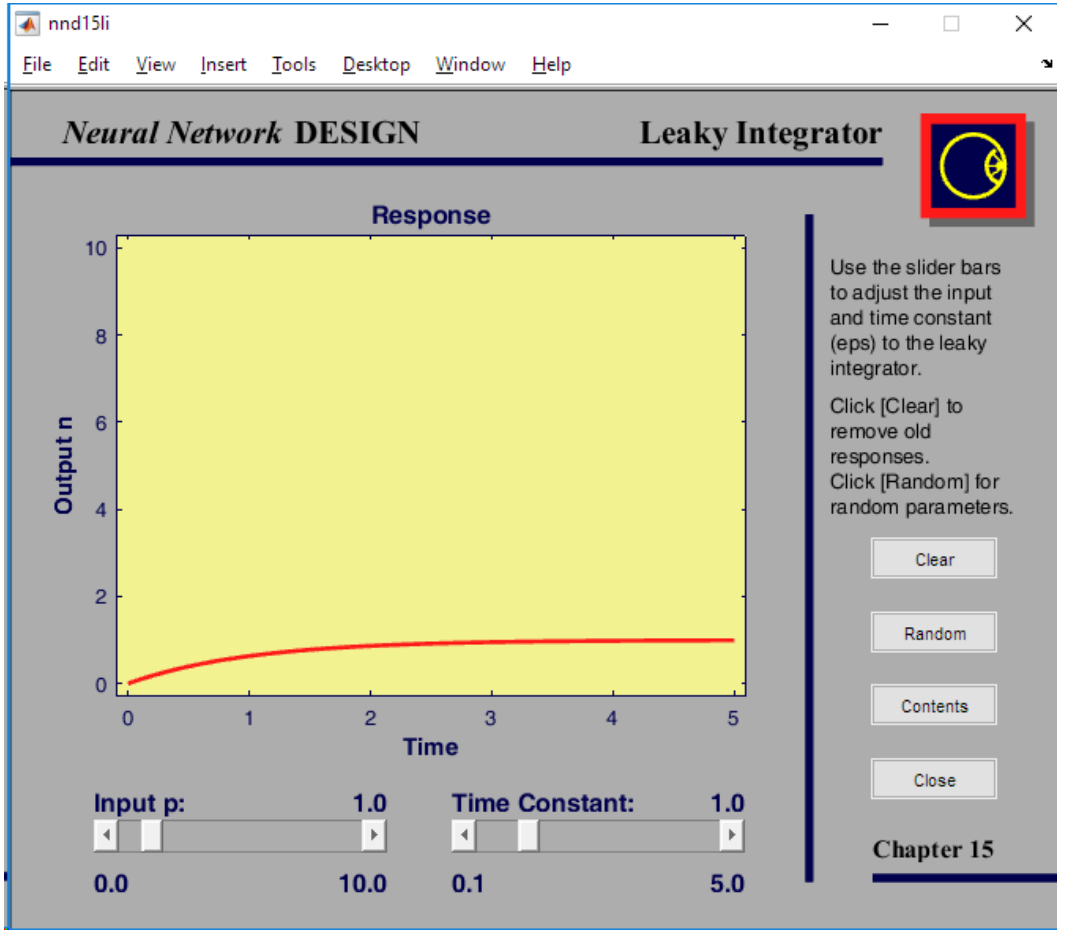
خصوصیات انتگرال‌گیری نشتی‌دار

## LEAKY INTEGRATOR



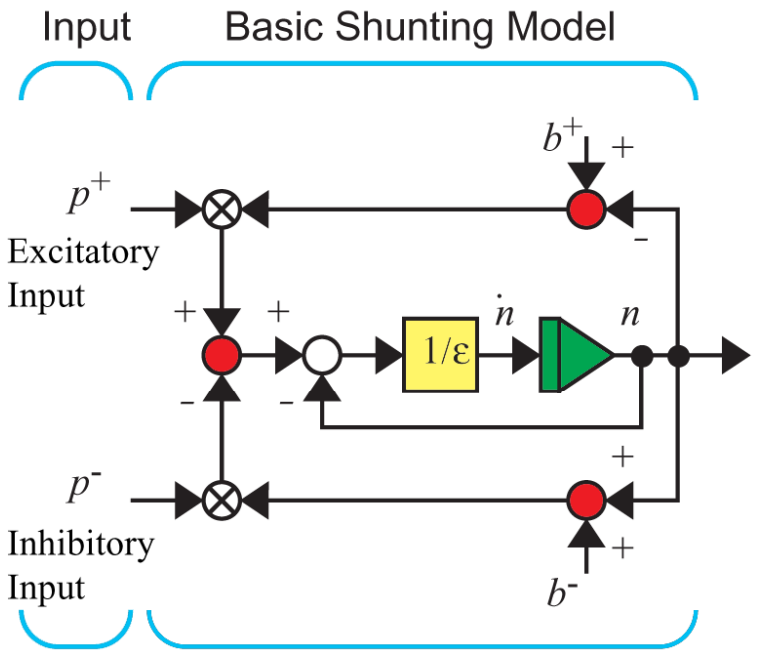

چون معادله‌ی پایه‌ی سیستم، خطی است،
در صورتی که ورودی $p$ مقیاس شود، آن‌گاه خروجی $n$ نیز به‌همان میزان مقیاس می‌شود.

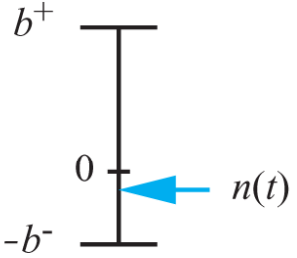
سرعت پاسخ با توجه به $\varepsilon$ (ثابت زمانی سیستم) مشخص می‌شود:
زمانی که $\varepsilon$ کاهش می‌یابد، پاسخ سریع‌تر می‌شود؛
زمانی که $\varepsilon$ افزایش می‌یابد، پاسخ آهسته‌تر می‌شود.


nnd15li
File Edit View Insert Tools Desktop Window Help
Neural Network DESIGN
Leaky Integrator
Response
Output n
Time
0 2 4 6 8 10
0 1 2 3 4 5
Use the slider bars to adjust the input and time constant (eps) to the leaky integrator.
Click [Clear] to remove old responses.
Click [Random] for random parameters.
Clear
Random
Contents
Close
Chapter 15
Input p:
1.0
0.0
10.0
Time Constant:
1.0
0.1
5.0

```
>> nnd15li
```

# Shunting Model




$$\varepsilon\, dn/dt = -n + (b^+ - n)p^+ - (n + b^-)p^-$$

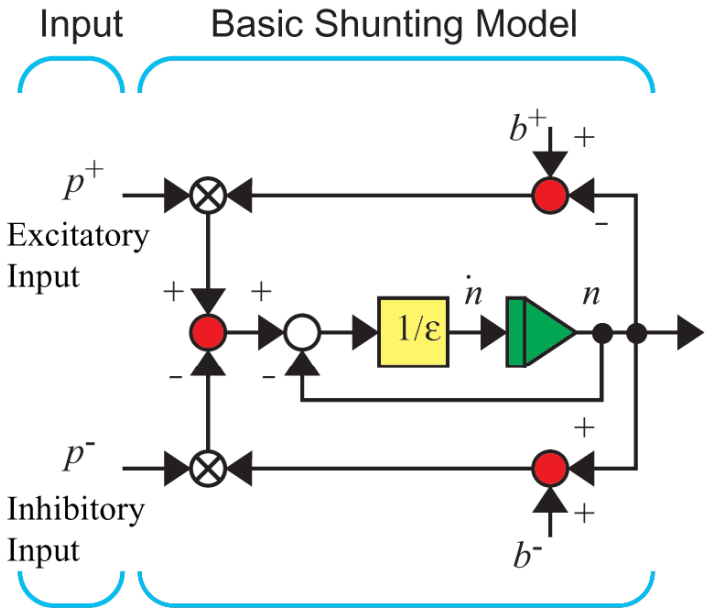
Gain Control (Sets upper limit)

Gain Control (Sets lower limit)

# مدل شانتینگ

زیربنای شبکه‌های رقابتی گراسبرگ

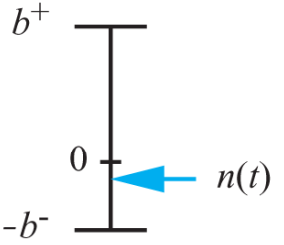
## SHUNTING MODEL




$$\varepsilon\, dn/dt = -n + (b^+ - n)p^+ - (n + b^-)p^-$$

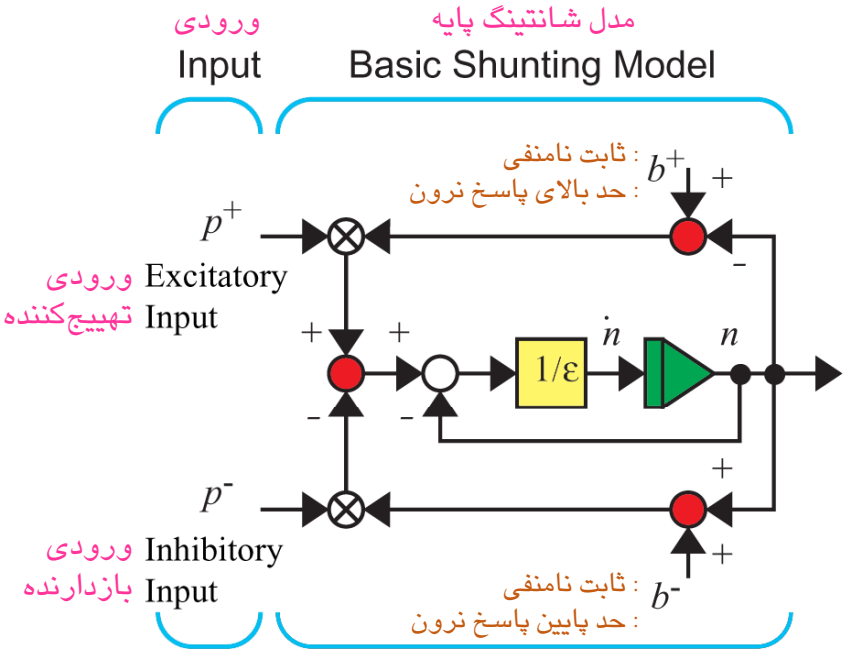
Gain Control (Sets upper limit)

Gain Control (Sets lower limit)

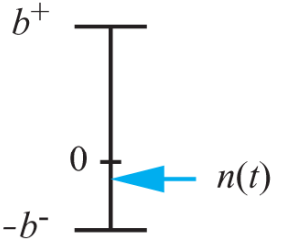


## مدل شانتینگ

SHUNTING MODEL




در صورتی که $n(0)$ بین $b^+$ و $-b^-$ باشد، آن‌گاه $n(t)$ بین این دو حد باقی می‌ماند.



# مدل شانتینگ

## معادله

SHUNTING MODEL

معادله دارای سه جمله در سمت راست است:

* اگر علامت این سه جمله مثبت باشد ⇐ $n(t)$ بالا خواهد رفت.
* اگر علامت این سه جمله منفی باشد ⇐ $n(t)$ پایین خواهد رفت.

$$\varepsilon \frac{dn(t)}{dt} = -n(t) + (b^+ - n(t))p^+ - (n(t) + b^-)p^-$$

جمله‌ی امحای خطی

جمله‌ی مهیا کننده‌ی کنترل بهره‌ی غیرخطی
(اگر $n(t)$ کمتر از $b^+$ باشد، مقدار این جمله مثبت خواهد بود.)
⇐ فراهم شدن حد بالا برای $n(t)$

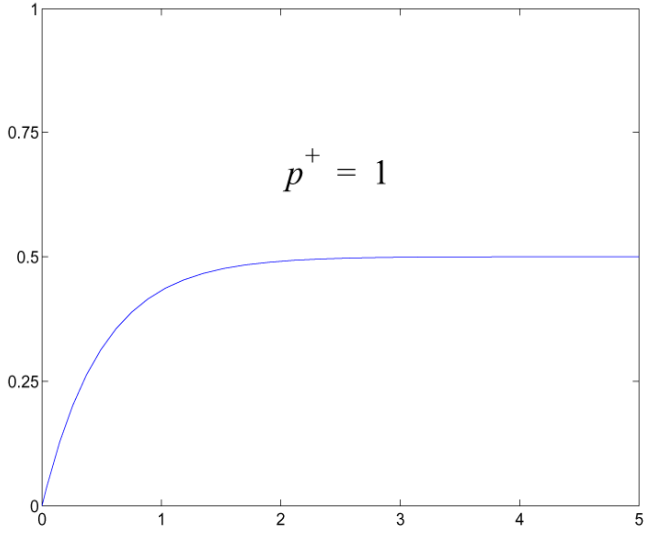
جمله‌ی مهیا کننده‌ی کنترل بهره‌ی غیرخطی
(اگر $n(t)$ بیشتر از $-b^-$ باشد، مقدار این جمله منفی خواهد بود.)
⇐ فراهم شدن حد پایین برای $n(t)$

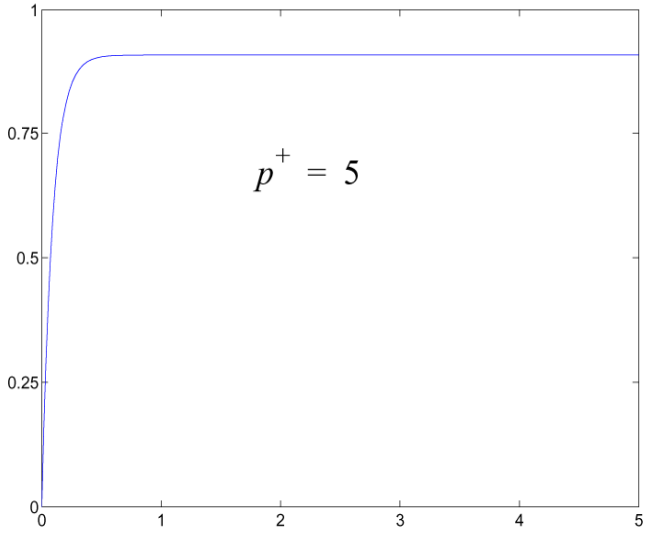
# Shunting Model Response

$$\varepsilon \frac{dn(t)}{dt} = -n(t) + (b^+ - n(t))p^+ - (n(t) + b^-)p^-$$

$$b^+ = 1 \qquad b^- = 0 \qquad \varepsilon = 1 \qquad p^- = 0$$

Upper limit will be 1, and lower limit will be 0.



# مدل شانتینگ

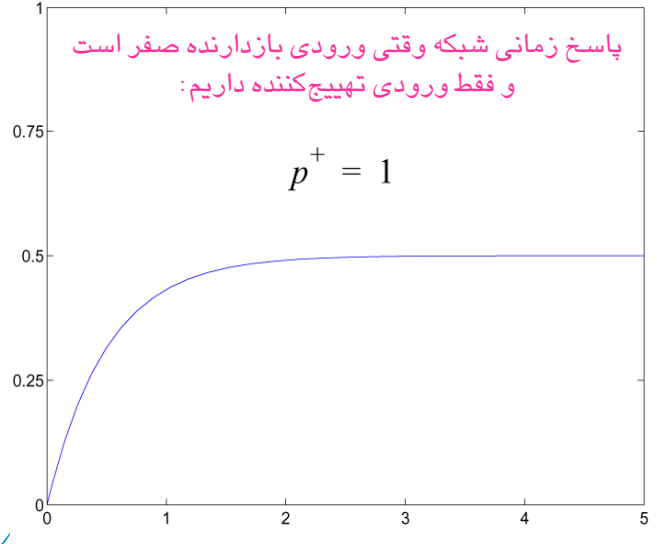
## پاسخ مدل شانتینگ

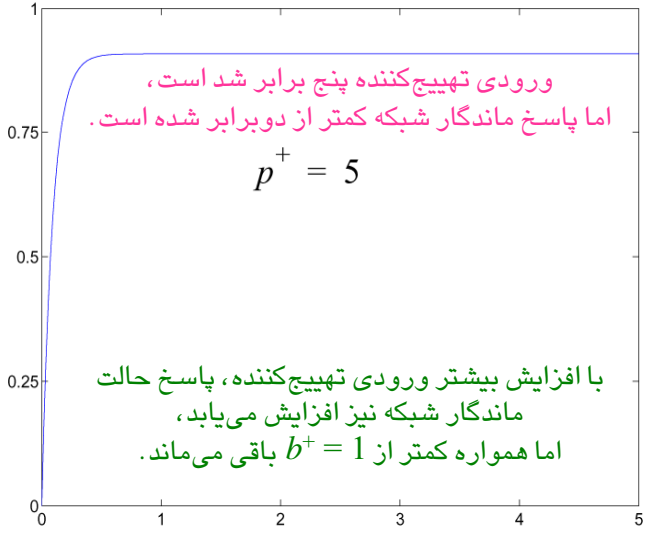
SHUNTING MODEL RESPONSE

$$\varepsilon \frac{dn(t)}{dt} = -n(t) + (b^+ - n(t))p^+ - (n(t) + b^-)p^-$$

$$b^+ = 1 \qquad b^- = 0 \qquad \varepsilon = 1 \qquad p^- = 0$$

Upper limit will be 1, and lower limit will be 0.




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

# مدل شانتینگ

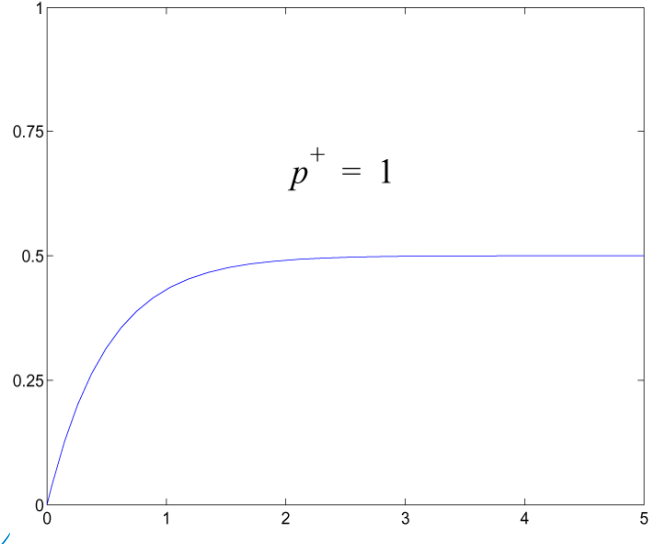
## پاسخ مدل شانتینگ

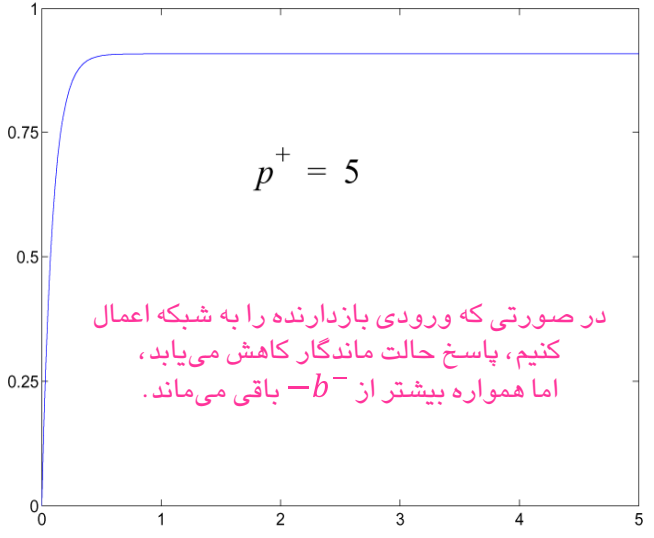
## SHUNTING MODEL RESPONSE

$$\varepsilon \frac{dn(t)}{dt} = -n(t) + (b^+ - n(t))p^+ - (n(t) + b^-)p^-$$

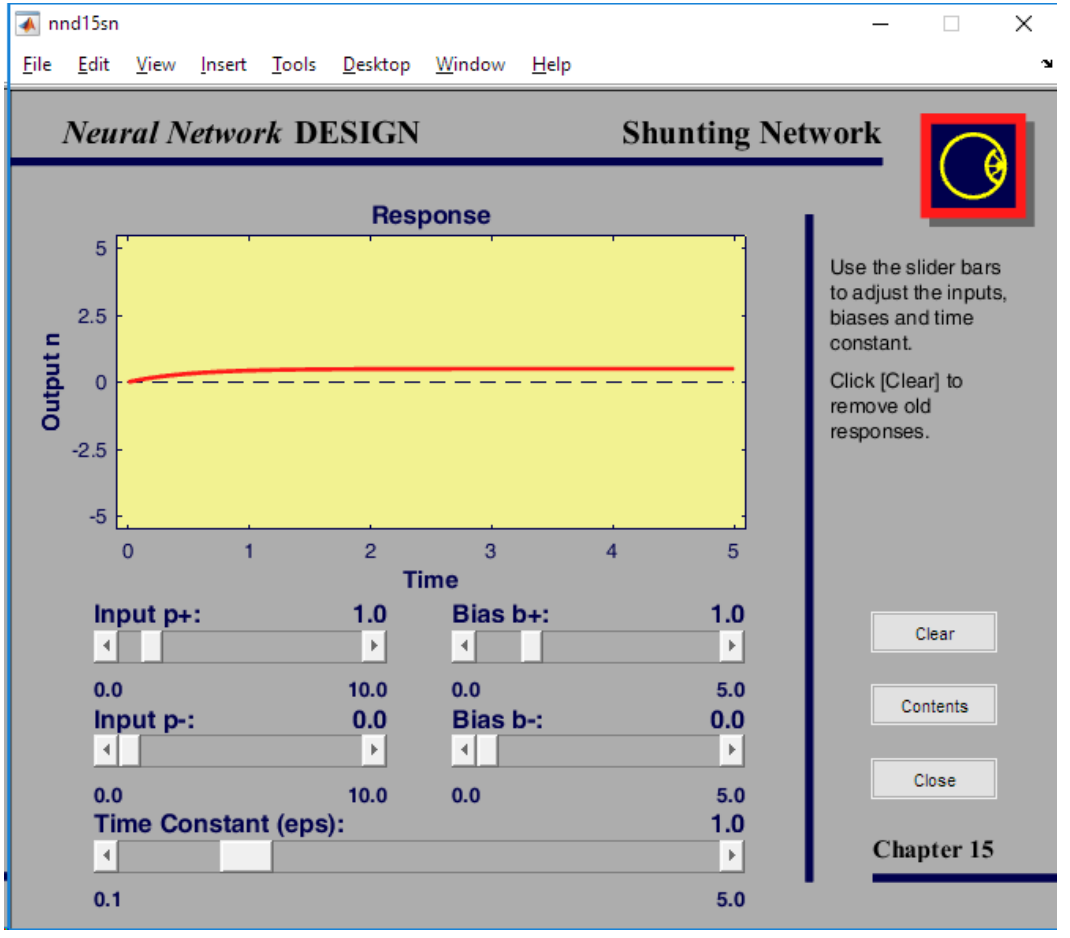
$$b^+ = 1 \qquad b^- = 0 \qquad \varepsilon = 1 \qquad p^- = 0$$

Upper limit will be 1, and lower limit will be 0.




در صورتی که ورودی بازدارنده را به شبکه اعمال کنیم، پاسخ حالت ماندگار کاهش می‌یابد، اما همواره بیشتر از $-b^-$ باقی می‌ماند.


nnd15sn
File Edit View Insert Tools Desktop Window Help
Neural Network DESIGN
Shunting Network
Response
Output n
5
2.5
0
-2.5
-5
0 1 2 3 4 5
Time
Input p+: 1.0
0.0 10.0
Bias b+: 1.0
0.0 5.0
Input p-: 0.0
0.0 10.0
Bias b-: 0.0
0.0 5.0
Time Constant (eps): 1.0
0.1 5.0
Use the slider bars to adjust the inputs, biases and time constant.
Click [Clear] to remove old responses.
Clear
Contents
Close
Chapter 15

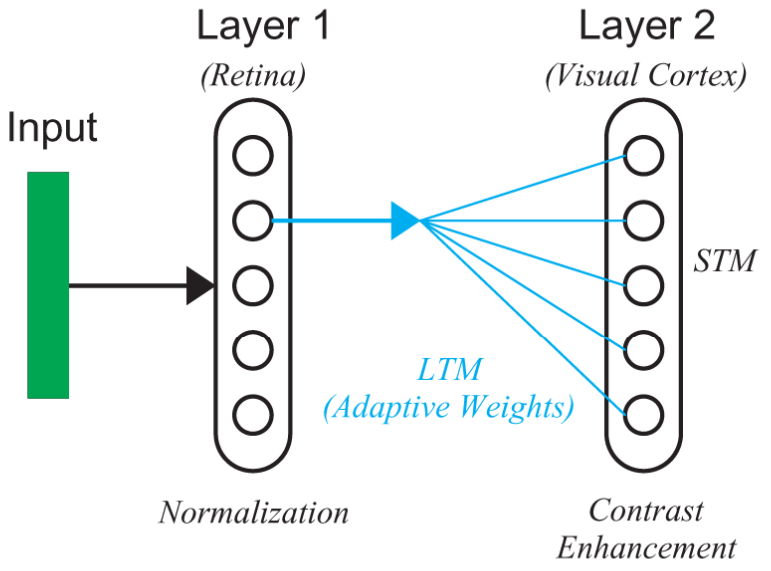
```
>> nnd15sn
```

شبکه‌ی گراسبرگ

# ۳

# شبکه‌ی رقابتی دولایه

# Grossberg Network

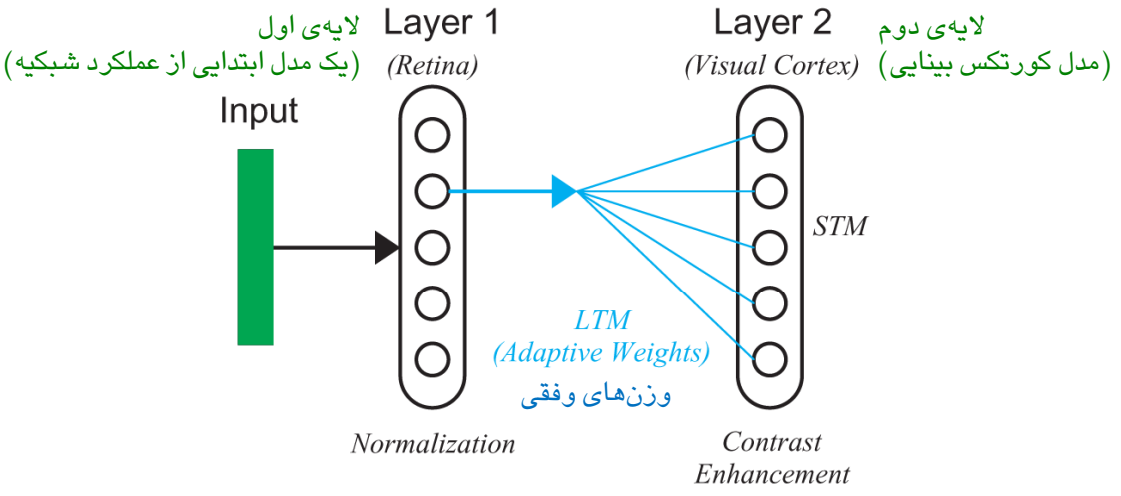



LTM - Long Term Memory (Network Weights)
STM - Short Term Memory (Network Outputs)

# شبکه‌ی گراسبرگ

شبکه‌ی رقابتی گراسبرگ (با توجه به نوع عملکرد سیستم بینایی انسان)

GROSSBERG NETWORK

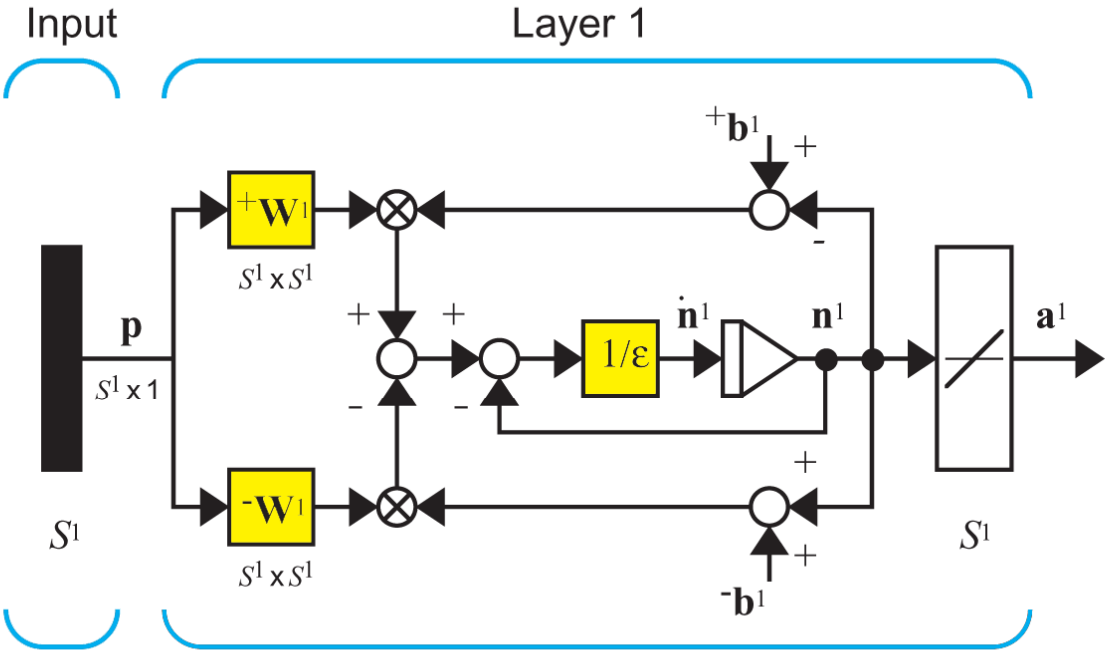



LTM - Long Term Memory (Network Weights)
STM - Short Term Memory (Network Outputs)

شبکه دارای دو مکانیسم LTM و STM است و اعمال
**وفق‌دهی**، **فیلتر کردن**، **نرمال‌سازی** و **ارتقای کنتراست** را انجام می‌دهد.

* این دو مدل قادر به پوشش کامل پیچیدگی‌های سیستم بینایی انسان نیستند،
اما از برخی خصوصیات آن بهره‌مند می‌باشند.

# Layer 1




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

# شبکه‌ی گراسبرگ

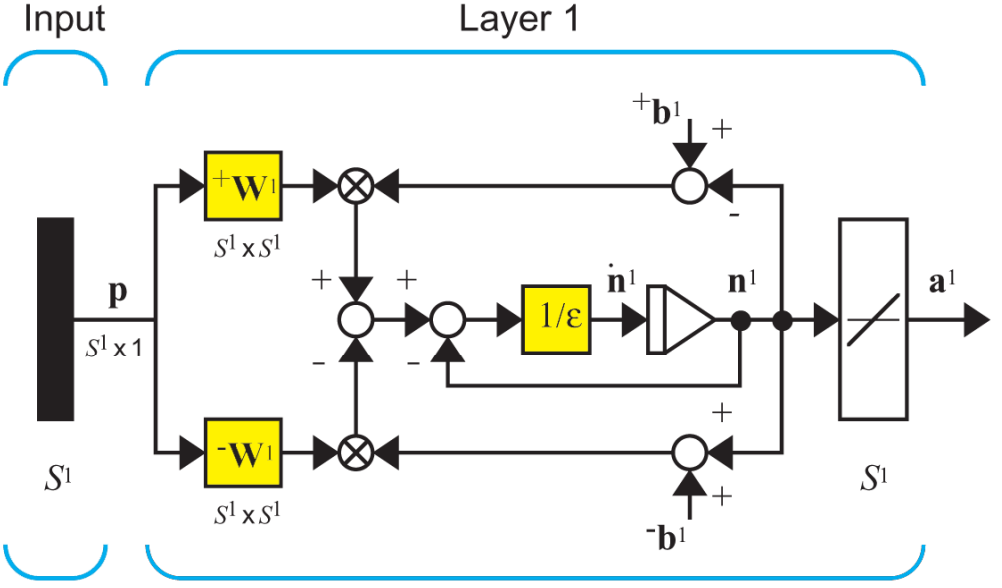
## لایه‌ی ۱

## LAYER 1

لایه‌ی ۱ : ورودی خارجی را دریافت کرده و شدت الگوی ورودی را نرمال می‌کند.

در ساختار این لایه از مدل شانتینگ استفاده شده است.

در این مدل ورودی‌های تهییج‌کننده و بازدارنده از روی بردار ورودی **p** محاسبه می‌شوند.




معادله‌ی عملکرد:

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

ثابت ε به‌گونه‌ای انتخاب می‌شود که عکس‌العمل نرون بسیار سریع‌تر از تغییر وزن‌های وفقی باشد.

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

# Operation of Layer 1

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

$${}^{-}\mathbf{b}^1 = \mathbf{0}$$

$${}^{+}b_i^1 = {}^{+}b^1$$

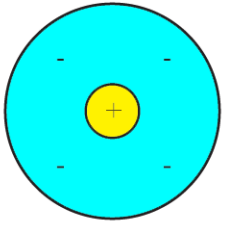
Excitatory Input

$$[{}^{+}\mathbf{W}^1]\mathbf{p} \qquad {}^{+}\mathbf{W}^1 = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

Inhibitory Input

$$[{}^{-}\mathbf{W}^1]\mathbf{p} \qquad {}^{-}\mathbf{W}^1 = \begin{bmatrix} 0 & 1 & \cdots & 1 \\ 1 & 0 & \cdots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \cdots & 0 \end{bmatrix}$$

On-Center/
Off-Surround
Connection
Pattern




Normalizes the input while maintaining relative intensities.

# شبکه‌ی گراسبرگ

## عملکرد لایه‌ی ۱

## OPERATION OF LAYER 1

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

$${}^{-}\mathbf{b}^1 = \mathbf{0}$$

$${}^{+}b_i^1 = {}^{+}b^1$$

Excitatory Input ورودی تهییج‌گر:

$[{}^{+}\mathbf{W}^1]\mathbf{p}$

ورودی تهییج‌گر نرون $i$–ام = $i$–امین عضو بردار ورودی است.

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

Inhibitory Input ورودی بازدارنده:

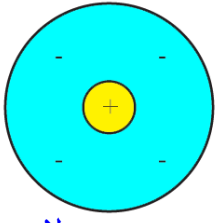
$[{}^{-}\mathbf{W}^1]\mathbf{p}$

ورودی بازدارنده‌ی نرون $i$–ام = مجموع همه‌ی اعضای بردار ورودی غیر از عضو $i$–ام

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

الگوی اتصال مرکز–روشن/ پیرامون–خاموش

On-Center/ Off-Surround Connection Pattern




لایه‌ی ۱، ورودی را **نرمال‌سازی** می‌کند در حالی که شدت‌های نسبی را نگه می‌دارد.

Normalizes the input while maintaining relative intensities.

ورودی تهییج‌گر نرون $i$–ام از عضوی از بردار ورودی به مرکزیت همان مکان ($t$) می‌آید، در حالی که ورودی بازدارنده از مکان‌های پیرامون آم می‌آید.

# شبکه‌ی گراسبرگ

## عملکرد لایه‌ی ۱

## OPERATION OF LAYER 1

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

$${}^{-}\mathbf{b}^1 = \mathbf{0}$$

$${}^{+}b_i^1 = {}^{+}b^1$$

برای سادگی، بایاس بازدارنده $(-b^-)$ را برابر با صفر قرار می‌دهیم.
⇐ کران پایین مدل شانتینگ، صفر می‌شود.

برای سادگی، بایاس تهییج‌گر را برای همه‌ی نرون‌ها مساوی هم قرار می‌دهیم: $(\forall i \;\; {}^{+}b_i^1 = {}^{+}b^1)$
⇐ کران بالای همه‌ی نرون‌ها یکسان می‌شود.

# Analysis of Normalization

Neuron *i* response:

$$\varepsilon \frac{dn_i^1(t)}{dt} = -n_i^1(t) + ({}^+b^1 - n_i^1(t))p_i - n_i^1(t)\sum_{j \neq i} p_j$$

At steady state:

$$0 = -n_i^1 + ({}^+b^1 - n_i^1)p_i - n_i^1\sum_{j \neq i} p_j \quad \Rightarrow \quad n_i^1 = \frac{{}^+b^1 p_i}{1 + \sum_{j=1}^{S^1} p_j}$$

Define relative intensity:

$$\bar{p}_i = \frac{p_i}{P} \qquad \text{where} \qquad P = \sum_{j=1}^{S^1} p_j$$

Steady state neuron activity:

$$n_i^1 = \left(\frac{{}^+b^1 P}{1+P}\right)\bar{p}_i \qquad \text{Total activity:} \quad \sum_{j=1}^{S^1} n_j^1 = \sum_{j=1}^{S^1} \left(\frac{{}^+b^1 P}{1+P}\right)\bar{p}_j = \left(\frac{{}^+b^1 P}{1+P}\right) \leq {}^+b^1$$

# شبکه‌ی گراسبرگ

## عملکرد لایه‌ی ۱ : تحلیل نرمال‌سازی

## ANALYSIS OF NORMALIZATION

Neuron $i$ response: پاسخ نرون در لایه‌ی ۱ :

$$\varepsilon \frac{dn_i^1(t)}{dt} = -n_i^1(t) + ({}^+b^1 - n_i^1(t))p_i - n_i^1(t)\sum_{j \neq i} p_j$$

At steady state: در حالت ماندگار مشتق صفر است:

$$0 = -n_i^1 + ({}^+b^1 - n_i^1)p_i - n_i^1\sum_{j \neq i} p_j \quad \Rightarrow \quad n_i^1 = \frac{{}^+b^1 p_i}{1 + \sum_{j=1}^{S^1} p_j}$$

Define relative intensity: مفهوم شدت نسبی ورودی $i$ را به صورت زیر تعریف می‌کنیم:

$$\bar{p}_i = \frac{p_i}{P} \quad \text{where} \quad P = \sum_{j=1}^{S^1} p_j$$

Steady state neuron activity: در این صورت، فعالیت نرون در حالت ماندگار می‌شود:

$$n_i^1 = \left(\frac{{}^+b^1 P}{1+P}\right)\bar{p}_i \qquad \text{Total activity:} \quad \sum_{j=1}^{S^1} n_j^1 = \sum_{j=1}^{S^1} \left(\frac{{}^+b^1 P}{1+P}\right)\bar{p}_j = \left(\frac{{}^+b^1 P}{1+P}\right) \leq {}^+b^1$$

پس $n_i^1$ متناسب خواهد بود با شدت نسبی $\bar{p}_i$ و از اندازه‌ی ورودی کلی $\mathbf{p}$ مستقل است.

# شبکه‌ی گراسبرگ

## عملکرد لایه‌ی ۱ : تحلیل نرمال‌سازی

ANALYSIS OF NORMALIZATION

$$n_i^1 = \left(\frac{{}^+b^1 P}{1+P}\right)\bar{p}_i \qquad \text{Total activity:} \quad \sum_{j=1}^{S^1} n_j^1 = \sum_{j=1}^{S^1} \left(\frac{{}^+b^1 P}{1+P}\right)\bar{p}_j = \left(\frac{{}^+b^1 P}{1+P}\right) \leq {}^+b^1$$

محدوده‌ی فعالیت نرون کران‌دار است:

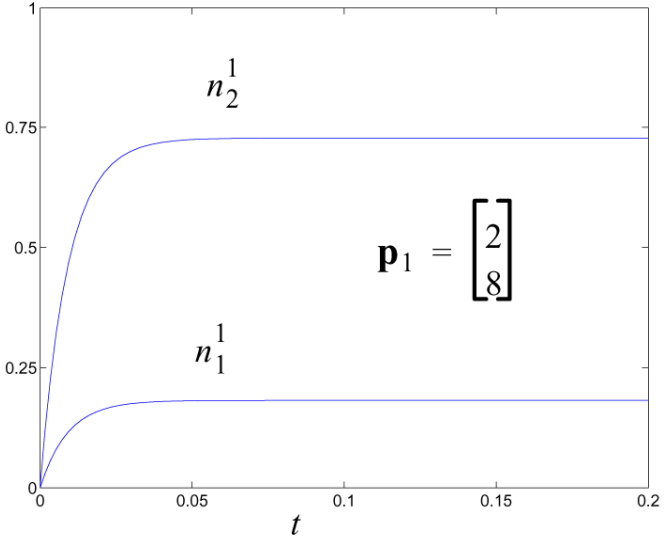
- بردار ورودی نرمال شده است، پس میزان فعالیت کل کمتر از ${}^+b^1$ خواهد بود.
  این در حالی است که شدت‌های نسبی هر یک از اعضای بردار ورودی برقرار شده است.
  ⇐ خروجی‌های لایه‌ی ۱ ($n_i^1$)، شدت‌های نسبی ورودی ($\bar{p}_i$) را علاوه بر نوسان‌های لحظه‌ای فعالیت ورودی **p** کد می‌کنند.
  این موضوع، نتیجه‌ی الگوی اتصال «مرکز–روشن/پیرامون–خاموش» ورودی‌ها و کنترل بهره‌ی غیرخطی مدل شانتینگ است.

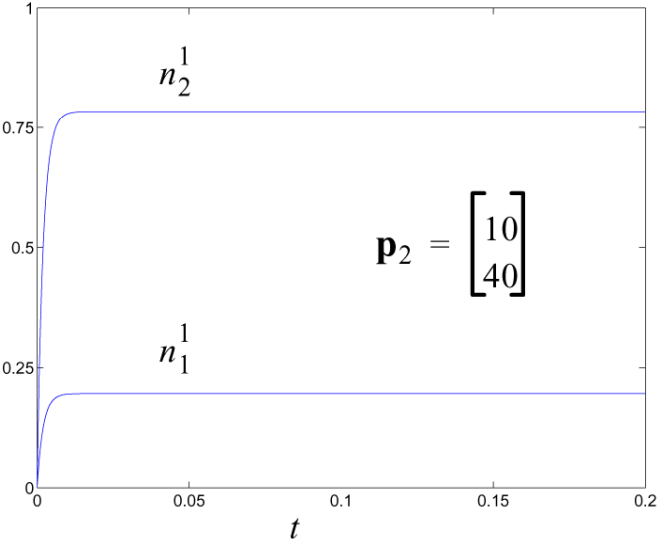
- لایه‌ی ۱ شبکه‌ی گراسبرگ، نوعی شبیه‌سازی از خصوصیات «ثبات روشنایی» و «کنتراست روشنایی» از سیستم بینایی انسان است:
  **شبکیه، به شدت نسبی تصویر، علاوه بر شدت مطلق آن حساس می‌باشد.**
  علاوه بر آن، آزمایش‌ها نشان می‌دهد که الگوی اتصال «مرکز–روشن/پیرامون–خاموش» یک ویژگی خاص از نواحی دریافت‌گر سلول‌های لنفاوی شبکیه است که دریافت‌گرهای نوری یک سلول را تغذیه می‌کند.

# Layer 1 Example

$$(0.1)\frac{dn_1^1(t)}{dt} = -n_1^1(t) + (1 - n_1^1(t))p_1 - n_1^1(t)p_2$$

$$(0.1)\frac{dn_2^1(t)}{dt} = -n_2^1(t) + (1 - n_2^1(t))p_2 - n_2^1(t)p_1$$



# شبکه‌ی گراسبرگ

## عملکرد لایه‌ی ۱ : مثال

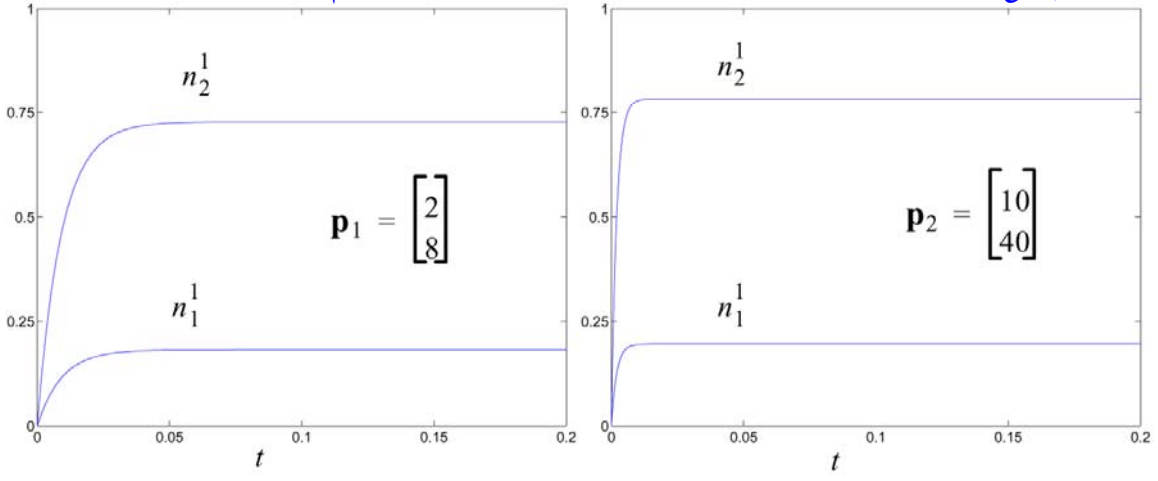
## LAYER 1 EXAMPLE

یک لایه با دو نرون با $b^1 = 1^+$ و $\varepsilon = 0.1$ در نظر می‌گیریم:

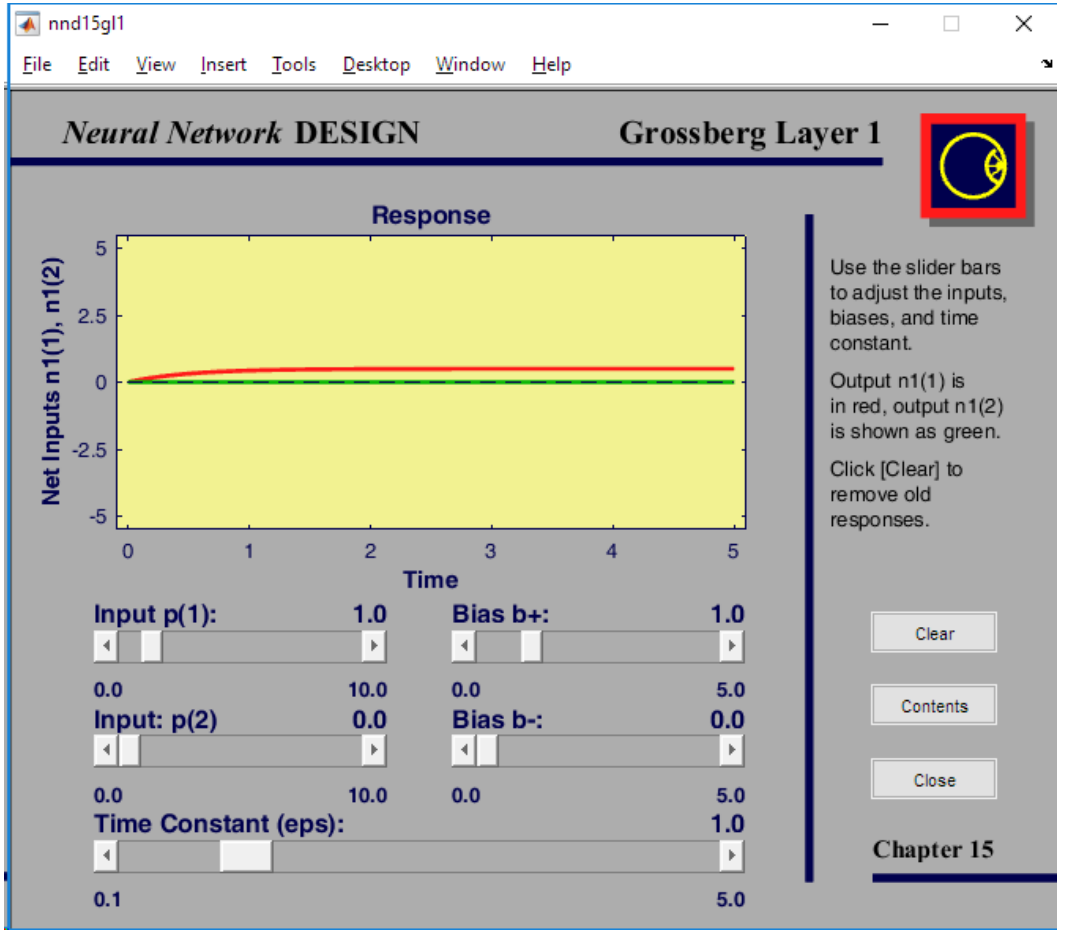
$$(0.1)\frac{dn_1^1(t)}{dt} = -n_1^1(t) + (1 - n_1^1(t))p_1 - n_1^1(t)p_2$$

$$(0.1)\frac{dn_2^1(t)}{dt} = -n_2^1(t) + (1 - n_2^1(t))p_2 - n_2^1(t)p_1$$

پاسخ شبکه برای دو بردار ورودی متفاوت: در هر دو بردار، عضو دوم، ۴ برابر عضو اول است:




⇐ پاسخ شبکه، ایجاد کننده‌ی شدت نسبی ورودی‌هاست؛ پاسخ کل $n_1^1(t) + n_2^1(t)$ همواره کمتر از 1 است.


nnd15gl1
File Edit View Insert Tools Desktop Window Help
Neural Network DESIGN
Grossberg Layer 1
Response
Net Inputs n1(1), n1(2)
5
2.5
0
-2.5
-5
0 1 2 3 4 5
Time
Input p(1): 1.0
0.0 10.0
Bias b+: 1.0
0.0 5.0
Input: p(2) 0.0
0.0 10.0
Bias b-: 0.0
0.0 5.0
Time Constant (eps): 1.0
0.1 5.0
Use the slider bars to adjust the inputs, biases, and time constant.
Output n1(1) is in red, output n1(2) is shown as green.
Click [Clear] to remove old responses.
Clear
Contents
Close
Chapter 15

```
>> nnd15gl1
```

# Characteristics of Layer 1

- The network is sensitive to relative intensities of the input pattern, rather than absolute intensities.
- The output of Layer 1 is a normalized version of the input pattern.
- The on-center/off-surround connection pattern and the nonlinear gain control of the shunting model produce the normalization effect.
- The operation of Layer 1 explains the brightness constancy and brightness contrast characteristics of the human visual system.

# شبکه‌ی گراسبرگ

## مشخصه‌های لایه ۱

## CHARACTERISTICS OF LAYER 1

- شبکه به شدت‌های نسبی الگوی ورودی حساس است، علاوه بر شدت‌های مطلق.
- خروجی لایه‌ی ۱ یک نسخه‌ی نرمال شده از الگوی ورودی است.
- الگوی اتصال مرکز–روشن/ پیرامون–خاموش و بهره‌ی کنترل غیرخطی مدل شانتینگ، اثر نرمال‌سازی را ایجاد می‌کنند.
- عملکرد لایه‌ی ۱ مشخصه‌های ثبات روشنایی و کنتراست روشنایی سیستم بینایی انسان را توضیح می‌دهد.

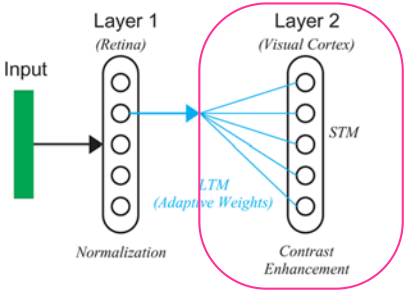
- The network is sensitive to relative intensities of the input pattern, rather than absolute intensities.
- The output of Layer 1 is a normalized version of the input pattern.
- The on-center/off-surround connection pattern and the nonlinear gain control of the shunting model produce the normalization effect.
- The operation of Layer 1 explains the brightness constancy and brightness contrast characteristics of the human visual system.

# شبکه‌ی گراسبرگ

## لایه‌ی ۲

## LAYER 2

لایه‌ی ۲ شبکه‌ی گراسبرگ: یک لایه از نرون‌های Instar پیوسته–زمان است
و چند کارکرد مختلف را پیاده‌سازی می‌کند:



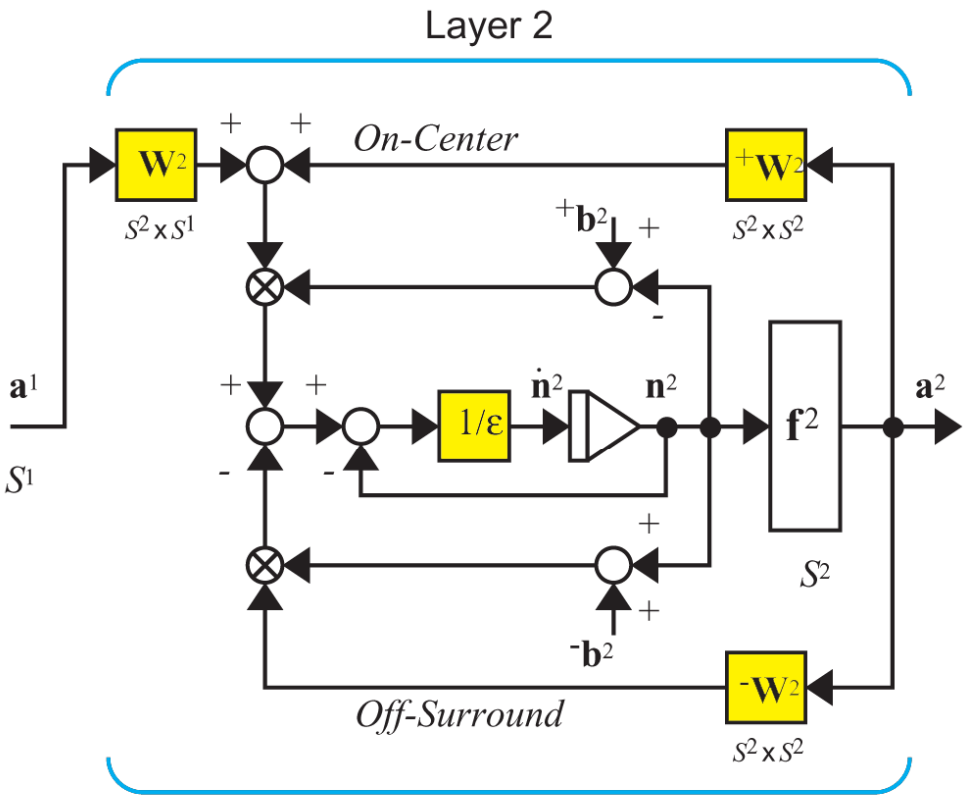

۱. مشابه لایه‌ی ۱، کل فعالیت لایه، نرمال‌سازی می‌شود.

۲. الگوهای مربوط به خود را ارتقای کنتراست می‌دهد
⇐ نرونی که بیشترین ورودی را دریافت می‌کند، بر پاسخ مسلط است.

۳. این لایه به‌عنوان یک حافظه‌ی کوتاه‌مدت STM عمل می‌کند و الگوهای ارتقای کنتراست یافته را ذخیره می‌کند.

تفاوت اصلی لایه‌ی ۲ با لایه‌ی ۱: **لایه‌ی ۲ از یک اتصال فیدبک استفاده می‌کند:**
این فیدبک:
۱) به شبکه اجازه‌ی ذخیره‌سازی یک الگو را می‌دهد (حتی در صورت حذف شدن ورودی).
۲) رقابتی را به‌انجام می‌رساند که منجر به ارتقای کنتراست می‌شود.

# Layer 2

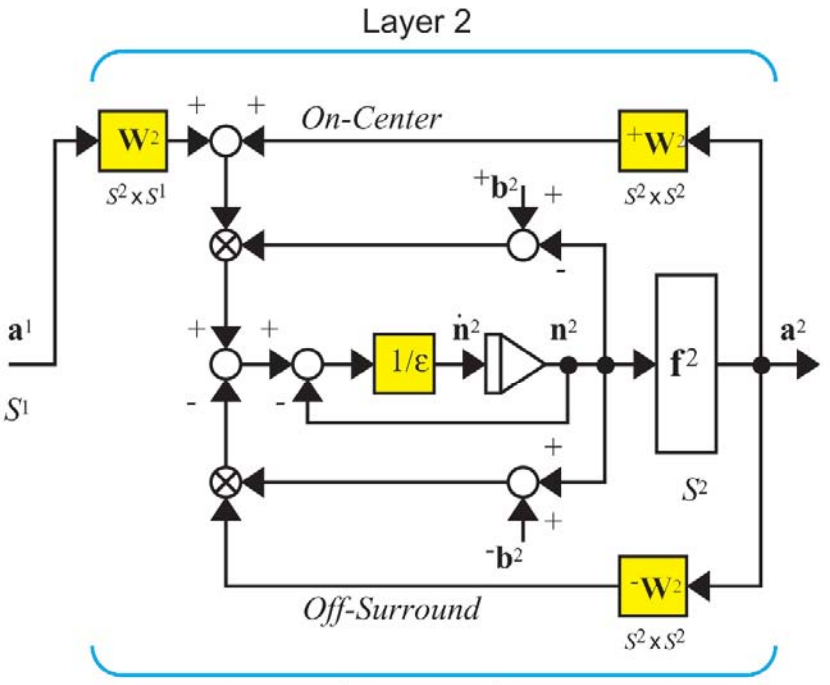



$$\varepsilon d\mathbf{n}^2/dt = -\mathbf{n}^2 + (^{+}\mathbf{b}^2 - \mathbf{n}^2)\{[^{+}\mathbf{W}^2]\mathbf{f}^2(\mathbf{n}^2) + \mathbf{W}^2\mathbf{a}^1\}$$
$$- (\mathbf{n}^2 + {}^{-}\mathbf{b}^2)[^{-}\mathbf{W}^2]\mathbf{f}^2(\mathbf{n}^2)$$

# شبکه‌ی گراسبرگ

لایه‌ی ۲

## LAYER 2




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

مدل شانتینگ پایه‌ی لایه‌ی دوم را تشکیل می‌دهد.

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

# Layer 2 Operation

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

Excitatory Input:

$$\{[{}^{+}\mathbf{W}^2]\mathbf{f}^2(\mathbf{n}^2(t)) + \mathbf{W}^2\mathbf{a}^1\}$$

$${}^{+}\mathbf{W}^2 = {}^{+}\mathbf{W}^1 \qquad \text{(On-center connections)}$$

$$\mathbf{W}^2 \qquad \text{(Adaptive weights)}$$

Inhibitory Input:

$$[{}^{-}\mathbf{W}^2]\mathbf{f}^2(\mathbf{n}^2(t))$$

$${}^{-}\mathbf{W}^2 = {}^{-}\mathbf{W}^1 \qquad \text{(Off-surround connections)}$$

# شبکه‌ی گراسبرگ

## عملکرد لایه‌ی ۲

## OPERATION OF LAYER 2

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

Excitatory Input: ورودی تهییج‌گر:

$$\{[{}^{+}\mathbf{W}^2]\mathbf{f}^2(\mathbf{n}^2(t)) + \mathbf{W}^2\mathbf{a}^1\}$$

ایجاد اتصالات فیدبک مرکز–روشن: $${}^{+}\mathbf{W}^2 = {}^{+}\mathbf{W}^1$$ (On-center connections)

وزن‌های وفقی (معادل با وزن‌های شبکه‌ی کوهونن): $$\mathbf{W}^2$$ (Adaptive weights)

Inhibitory Input: ورودی بازدارنده:

$$[{}^{-}\mathbf{W}^2]\mathbf{f}^2(\mathbf{n}^2(t))$$

ایجاد اتصالات فیدبک پیرامون–خاموش: $${}^{-}\mathbf{W}^2 = {}^{-}\mathbf{W}^1$$ (Off-surround connections)

# شبکه‌ی گراسبرگ

## عملکرد لایه‌ی ۲

## OPERATION OF LAYER 2

ورودی‌های لایه‌ی دوم از ضرب داخلی میان الگوهای اولیه (سطرهای ماتریس وزن $\mathbf{W}^2$)
و خروجی لایه‌ی اول (الگوهای ورودی نرمال‌شده) حاصل می‌شود.

بزرگ‌ترین ضرب داخلی، متناظر با نزدیک‌ترین الگوی اولیه با الگوی ورودی خواهد بود.

سپس لایه‌ی دوم یک رقابت میان نرون‌ها برگزار می‌کند که منجر به ارتقای کنتراست الگوی خروجی می‌شود
⇐ برجسته شدن خروجی‌های بزرگ و محو شدن الگوهای کوچک

* این نوع ارتقای کنتراست عموماً نسبت به رقابت «برنده همه را می‌خورد» در شبکه‌های همینگ و کوهونن ملایم‌تر است.
در آنها خروجی همه‌ی نرون‌های بازنده صفر است و فقط خروجی نرون برنده برابر با 1 است.
در شبکه‌ی گراسبرگ رقابت منجر به برجسته شدن مقادیر بزرگ‌تر و محو شدن مقادیر کوچک‌تر می‌شود
ولی خروجی مقادیر کوچک‌تر لزوماً برابر با صفر نیست.

مقدار ارتقای کنتراست در لایه‌ی دوم با توجه به تابع انتقال $f^2$ تعیین می‌شود.

# Layer 2 Example

$$\varepsilon = 0.1 \quad {}^{+}\mathbf{b}^2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad {}^{-}\mathbf{b}^2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad f^2(n) = \frac{10(n)^2}{1+(n)^2} \quad \mathbf{W}^2 = \begin{bmatrix} ({}_1\mathbf{w}^2)^T \\ ({}_2\mathbf{w}^2)^T \end{bmatrix} = \begin{bmatrix} 0.9 & 0.45 \\ 0.45 & 0.9 \end{bmatrix}$$

Correlation between prototype 1 and input.

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

Correlation between prototype 2 and input.

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

# شبکه‌ی گراسبرگ

## عملکرد لایه‌ی ۲ : مثال

### OPERATION OF LAYER 2

$$\varepsilon = 0.1 \quad {}^{+}\mathbf{b}^2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad {}^{-}\mathbf{b}^2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad f^2(n) = \frac{10(n)^2}{1+(n)^2} \quad \mathbf{W}^2 = \begin{bmatrix} ({}_1\mathbf{w}^2)^T \\ ({}_2\mathbf{w}^2)^T \end{bmatrix} = \begin{bmatrix} 0.9 & 0.45 \\ 0.45 & 0.9 \end{bmatrix}$$

Correlation between prototype 1 and input.

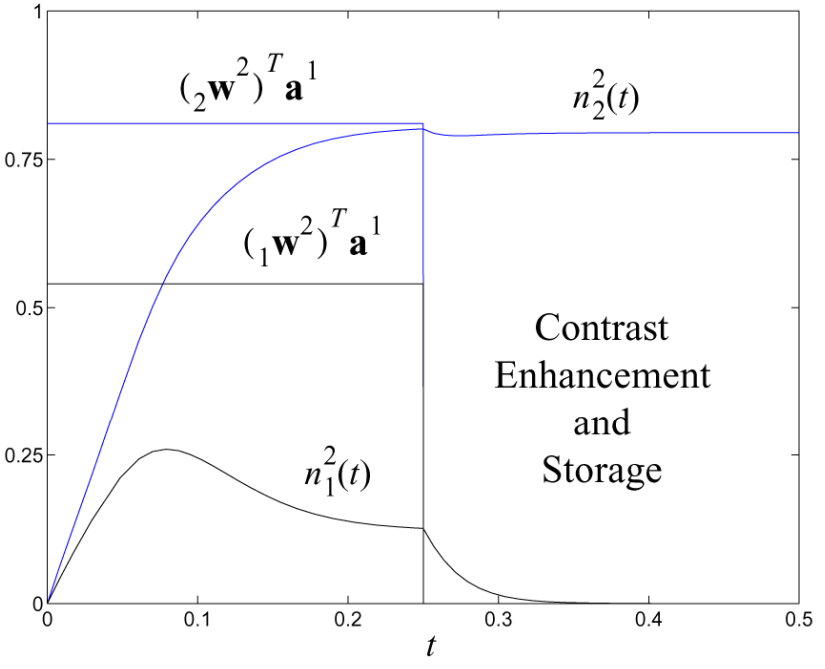
$$(0.1)\frac{dn_1^2(t)}{dt} = -n_1^2(t) + (1-n_1^2(t))\left\{ f^2(n_1^2(t)) + \overbrace{({}_1\mathbf{w}^2)^T\mathbf{a}^1} \right\} - n_1^2(t) f^2(n_2^2(t))$$

Correlation between prototype 2 and input.

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

# Layer 2 Response

$$\mathbf{a}^1 = \begin{bmatrix} 0.2 \\ 0.8 \end{bmatrix}$$




Input to neuron 1:

$$({}_1\mathbf{w}^2)^T\mathbf{a}^1 = \begin{bmatrix} 0.9 & 0.45 \end{bmatrix}\begin{bmatrix} 0.2 \\ 0.8 \end{bmatrix} = 0.54$$

Input to neuron 2:

$$({}_2\mathbf{w}^2)^T\mathbf{a}^1 = \begin{bmatrix} 0.45 & 0.9 \end{bmatrix}\begin{bmatrix} 0.2 \\ 0.8 \end{bmatrix} = 0.81$$

# شبکه‌ی گراسبرگ

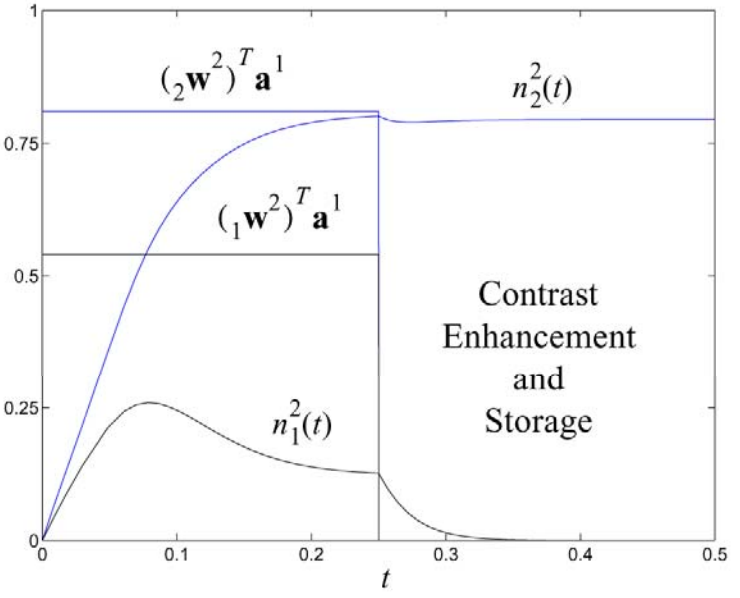
## مثال: پاسخ لایه‌ی ۲

### LAYER 2 RESPONSE

ورودی:

$$\mathbf{a}^1 = \begin{bmatrix} 0.2 \\ 0.8 \end{bmatrix}$$

نتیجه‌ی حالت ماندگار لایه‌ی ۱ به‌مدت 0.25 ثانیه به شبکه اعمال شده است و سپس حذف می‌شود.




دو خصوصیت مهم این پاسخ:

۱) حتی پیش از حذف ورودی، نوعی ارتقای کنتراست اتفاق می‌افتد.

۲) پس از اینکه ورودی برابر با صفر قرار داده می‌شود، شبکه میزان کنتراست را ارتقا داده و الگو را ذخیره می‌کند.
(به این رفتار شبکه، طنین – reverberation می‌گویند)

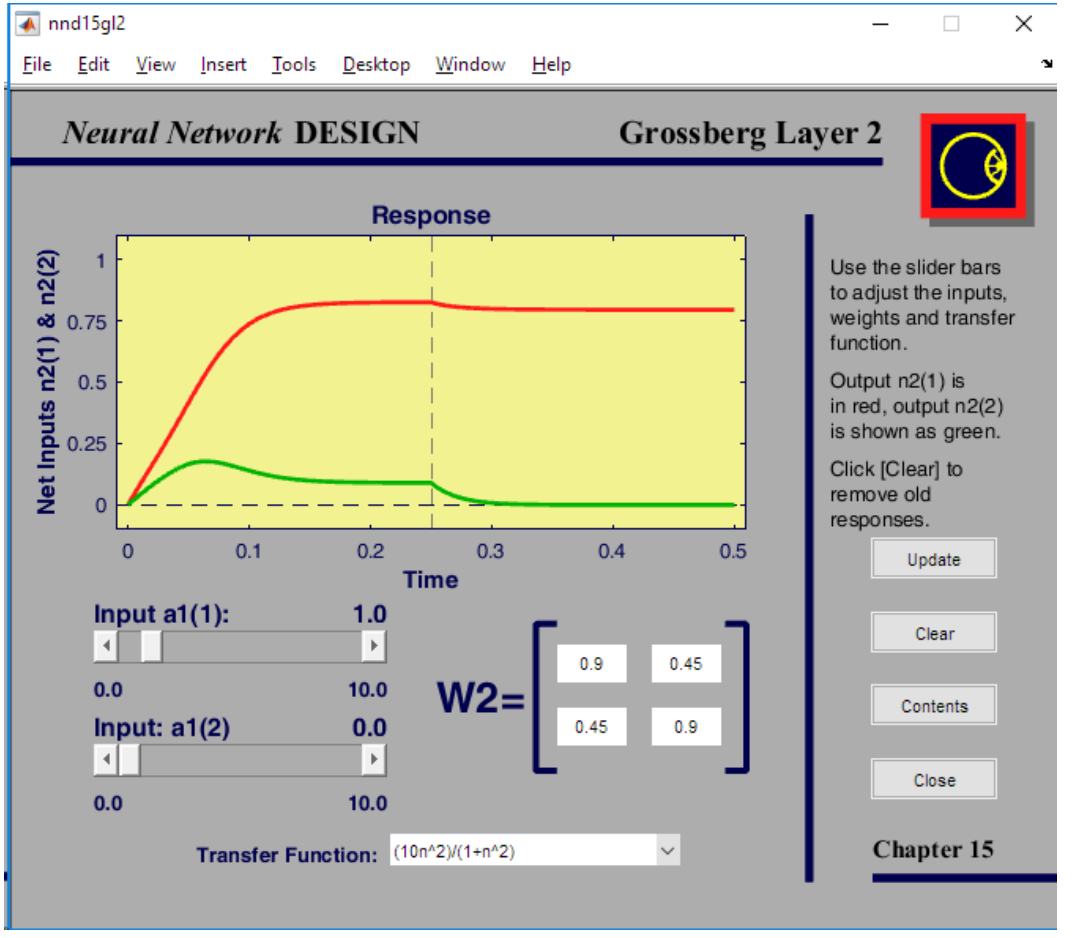
Input to neuron 1:

$$({}_1\mathbf{w}^2)^T\mathbf{a}^1 = \begin{bmatrix} 0.9 & 0.45 \end{bmatrix}\begin{bmatrix} 0.2 \\ 0.8 \end{bmatrix} = 0.54$$

Input to neuron 2:

$$({}_2\mathbf{w}^2)^T\mathbf{a}^1 = \begin{bmatrix} 0.45 & 0.9 \end{bmatrix}\begin{bmatrix} 0.2 \\ 0.8 \end{bmatrix} = 0.81$$

* **فیدبک غیر خطی** شبکه را قادر به ذخیره‌سازی الگو می‌کند.
* **اتصالات «مرکز–روشن/پیرامون–خاموش»** باعث ارتقای کنتراست می‌شوند.


nnd15gl2
File Edit View Insert Tools Desktop Window Help
Neural Network DESIGN
Grossberg Layer 2
Response
Net Inputs n2(1) & n2(2)
1
0.75
0.5
0.25
0
0 0.1 0.2 0.3 0.4 0.5
Time
Input a1(1): 1.0
0.0 10.0
Input: a1(2) 0.0
0.0 10.0
W2=
0.9 0.45
0.45 0.9
Transfer Function: (10n^2)/(1+n^2)
Use the slider bars to adjust the inputs, weights and transfer function.
Output n2(1) is in red, output n2(2) is shown as green.
Click [Clear] to remove old responses.
Update
Clear
Contents
Close
Chapter 15

```
>> nnd15gl2
```

# Characteristics of Layer 2

- As in the Hamming and Kohonen networks, the inputs to Layer 2 are the inner products between the prototype patterns (rows of the weight matrix $\mathbf{W}^2$) and the output of Layer 1 (normalized input pattern).
- The nonlinear feedback enables the network to store the output pattern (pattern remains after input is removed).
- The on-center/off-surround connection pattern causes contrast enhancement (large inputs are maintained, while small inputs are attenuated).

# شبکه‌ی گراسبرگ

## مشخصه‌های لایه ۲

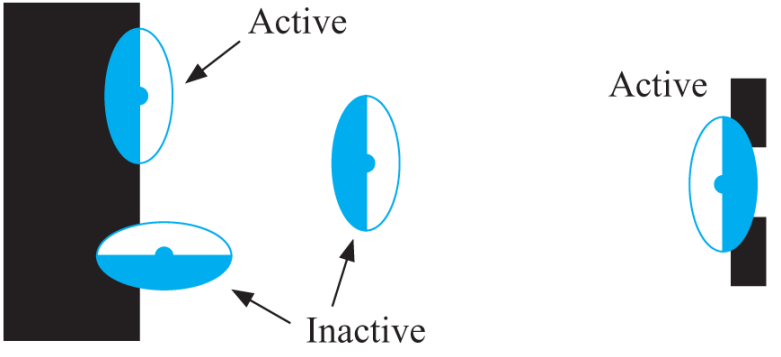
## CHARACTERISTICS OF LAYER 1

- As in the Hamming and Kohonen networks, the inputs to Layer 2 are the inner products between the prototype patterns (rows of the weight matrix $\mathbf{W}^2$) and the output of Layer 1 (normalized input pattern).
- The nonlinear feedback enables the network to store the output pattern (pattern remains after input is removed).
- The on-center/off-surround connection pattern causes contrast enhancement (large inputs are maintained, while small inputs are attenuated).

- همانند شبکه‌های همینگ و کوهونن، ورودی‌های لایه‌ی ۲ ضرب‌های داخلی میان الگوهای پروتوتایپ (سطرهای ماتریس وزن $\mathbf{W}^2$) و خروجی لایه‌ی ۱ (الگوی ورودی نرمال‌شده است).
- فیدبک غیرخطی شبکه را قادر می‌کند که الگوی خروجی را ذخیره کند (الگو پس از حذف شدن ورودی باقی می‌ماند).
- الگوی اتصالات «مرکز–روشن/پیرامون–خاموش» موجب ارتقای کنتراست می‌شود (ورودی‌های بزرگ حفظ می‌شوند، در حالی که ورودی‌های کوچک تضعیف می‌شوند.).

# Oriented Receptive Field

When an oriented receptive field is used, instead of an on-center/off-surround receptive field, the emergent segmentation problem can be understood.



# شبکه‌ی گراسبرگ

## میدان پذیرنده‌ی گرادار

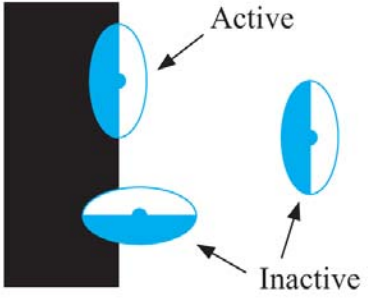
## ORIENTED RECEPTIVE FIELD

مسئله‌ی Emergent Segmentation:

از ساختار «میدان پذیرنده‌ی گرادار» برای پیاده‌سازی این مکانیسم استفاده شده است.

اتصالات بازدارنده (off) در سمت دیگر: : اتصالات تهییج‌گر (on) در یک سمت

When an oriented receptive field is used, instead of an on-center/off-surround receptive field, the emergent segmentation problem can be understood.




زمانی که میدان بر یک لبه منطبق باشد، میدان متناظر فعال خواهد شد و زمانی که میدان منطبق بر یک لبه نباشد، نرون غیرفعال خواهد بود.

این مثال نشان می‌دهد که چرا ممکن است در ناحیه‌ای که لبه وجود ندارد، ما یک لبه تشخیص دهیم:

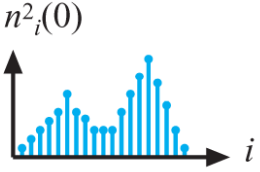
# شبکه‌ی گراسبرگ

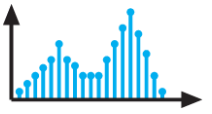
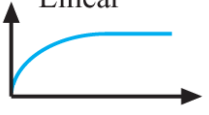
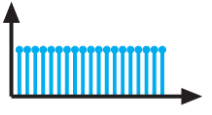
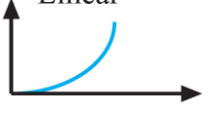
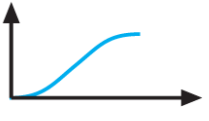
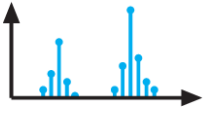
## انتخاب تابع انتقال لایه‌ی دوم

## CHOICE OF TRANSFER FUNCTION

رفتار لایه‌ی دوم شبکه‌ی گراسبرگ تا حد زیادی به تابع انتقال $f^2$ بستگی دارد.

# Choice of Transfer Function




| $f^2(n)$ | Stored Pattern $\mathbf{n}^2(\infty)$ | Comments |
|---|---|---|
| Linear |  | Perfect storage of any pattern, but amplifies noise. |
| Slower than Linear  |  | Amplifies noise, reduces contrast. |
| Faster than Linear  | | Winner-take-all, suppresses noise, quantizes total activity. |
| Sigmoid  |  | Supresses noise, contrast enhances, not quantized. |

# شبکه‌ی گراسبرگ

## انتخاب تابع انتقال لایه‌ی دوم

## CHOICE OF TRANSFER FUNCTION

نحوه‌ی تأثیر انتخاب $f^2(n)$ بر روی پاسخ حالت ماندگار شبکه در صورت حذف ورودی

برای مثال فرض کنید یک ورودی برای یک مدت زمان کوتاه به شبکه اعمال شود ⇐ خروجی مطابق شکل به حالت ماندگار می‌رسد.

| $f^2(n)$ | Stored Pattern $\mathbf{n}^2(\infty)$ | Comments |
|---|---|---|
| Linear | | Perfect storage of any pattern, but amplifies noise. |
| Slower than Linear | | Amplifies noise, reduces contrast. |
| Faster than Linear | | Winner-take-all, suppresses noise, quantizes total activity. |
| Sigmoid | | Supresses noise, contrast enhances, not quantized. |

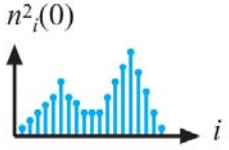
**تابع انتقال خطی:**
الگو به طور کامل ذخیره می‌شود، نویز همراه با الگو تقویت و در کنار آن ذخیره می‌شود.

**تابع انتقال آهسته‌تر از خطی:**
پاسخ حالت ماندگار مستقل از شرایط آغازین؛ همه‌ی کنتراست‌ها حذف و نویز تقویت می‌شود.

**تابع انتقال سریع‌تر از خطی:**
ایجاد رقابت «برنده همه را می‌خورد» نرون‌های دارای بیشترین مقدار آغازین ذخیره شده و بقیه حذف می‌شوند
⇐ حداقل شدن تأثیرات منفی نویز

**تابع انتقال سیگموئید:** برای سیگنال‌های کوچک: سریع‌تر از خطی؛
سیگنال‌های متوسط: تقریباً خطی؛
سیگنال‌های بزرگ: آهسته‌تر از خطی؛

هر نقطه نمایشگر یک نرون است:



# شبکه‌ی گراسبرگ

## انتخاب تابع انتقال لایه‌ی دوم: تابع سیگموئید

## CHOICE OF TRANSFER FUNCTION

زمانی که از تابع انتقال سیگموئید به‌عنوان $f^2(n)$ استفاده می‌شود، کنتراست الگو ارتقا داده می‌شود:

مقادیر بزرگ‌تر تقویت شده و مقادیر کوچک‌تر محو می‌شوند
⇐ خروجی‌هایی که کمتر از حد معینی هستند (آستانه‌ی اطفاء)، به سمت صفر محو می‌شوند.

⇐ خصوصیت حذف نویز در توابع انتقال سریع‌تر از خطی،
با خصوصیت ذخیره‌سازی کامل در مورد توابع انتقال خطی، ادغام می‌شود.

# Adaptive Weights

## Hebb Rule with Decay

$$\frac{dw^2_{i,j}(t)}{dt} = \alpha\{-w^2_{i,j}(t) + n^2_i(t)n^1_j(t)\}$$

## Instar Rule (Gated Learning)

$$\frac{dw^2_{i,j}(t)}{dt} = \alpha n^2_i(t)\{-w^2_{i,j}(t) + n^1_j(t)\}$$

Learn when $n_i^2(t)$ is active.

## Vector Instar Rule

$$\frac{d[{}_i\mathbf{w}^2(t)]}{dt} = \alpha n^2_i(t)\{-[{}_i\mathbf{w}^2(t)] + \mathbf{n}^1(t)\}$$

# شبکه‌ی گراسبرگ

## وزن‌های وفقی

### ADAPTIVE WEIGHTS

مانند شبکه‌های همینگ و کوهونن، الگوهای ذخیره شده‌ای که در نزدیک‌ترین حالت نسبت به الگوهای ورودی هستند، بزرگ‌ترین خروجی لایه‌ی دوم را تولید می‌کنند.

قانون یادگیری وزن‌های وفقی ($\mathbf{W}^2$):

Hebb Rule with Decay

$$\frac{dw^2_{i,j}(t)}{dt} = \alpha\{-w^2_{i,j}(t) + n^2_i(t)n^1_j(t)\}$$

جمله‌ی passive decay

گراسبرگ به این وزن‌های وفقی، **حافظه‌ی بلندمدت (LTM)** می‌گوید:

زیرا سطرهای $\mathbf{W}^2$ بازنمایی کننده‌ی الگوهای ذخیره‌شده‌ای است که شبکه توانایی تشخیص آنها را دارد.

Instar Rule
(Gated Learning)

$$\frac{dw^2_{i,j}(t)}{dt} = \alpha n^2_i(t)\{-w^2_{i,j}(t) + n^1_j(t)\}$$

Learn when $n_i^2(t)$ is active.

اغلب، غیر فعال کردن یادگیری (فراموشی) در صورت غیر فعال بودن $n_i^2(t)$ مفید است.

Vector Instar Rule

$$\frac{d[{}_i\mathbf{w}^2(t)]}{dt} = \alpha n^2_i(t)\{-[{}_i\mathbf{w}^2(t)] + \mathbf{n}^1(t)\}$$

طرف راست قاعده‌ی یادگیری در $n_i^2(t)$ ضرب شده است ⇐ یادگیری فقط در زمان غیر صفر بودن آن انجام می‌شود.
[پیاده‌سازی پیوسته–زمان قاعده‌ی یادگیری Instar]

# Example

$$\frac{dw^2_{1,1}(t)}{dt} = n^2_1(t)\{-w^2_{1,1}(t) + n^1_1(t)\}$$

$$\frac{dw^2_{1,2}(t)}{dt} = n^2_1(t)\{-w^2_{1,2}(t) + n^1_2(t)\}$$

$$\frac{dw^2_{2,1}(t)}{dt} = n^2_2(t)\{-w^2_{2,1}(t) + n^1_1(t)\}$$

$$\frac{dw^2_{2,2}(t)}{dt} = n^2_2(t)\{-w^2_{2,2}(t) + n^1_2(t)\}$$

# شبکه‌ی گراسبرگ

## وزن‌های وفقی: مثال

ADAPTIVE WEIGHTS

یک شبکه‌ی گراسبرگ با دو نرون در هر لایه:

معادلات به‌روزرسانی وزن‌ها:

$$\frac{dw^2_{1,1}(t)}{dt} = n^2_1(t)\{-w^2_{1,1}(t) + n^1_1(t)\}$$

$$\frac{dw^2_{1,2}(t)}{dt} = n^2_1(t)\{-w^2_{1,2}(t) + n^1_2(t)\}$$

$$\frac{dw^2_{2,1}(t)}{dt} = n^2_2(t)\{-w^2_{2,1}(t) + n^1_1(t)\}$$

$$\frac{dw^2_{2,2}(t)}{dt} = n^2_2(t)\{-w^2_{2,2}(t) + n^1_2(t)\}$$

نرخ یادگیری $\alpha = 1$

# Response of Adaptive Weights

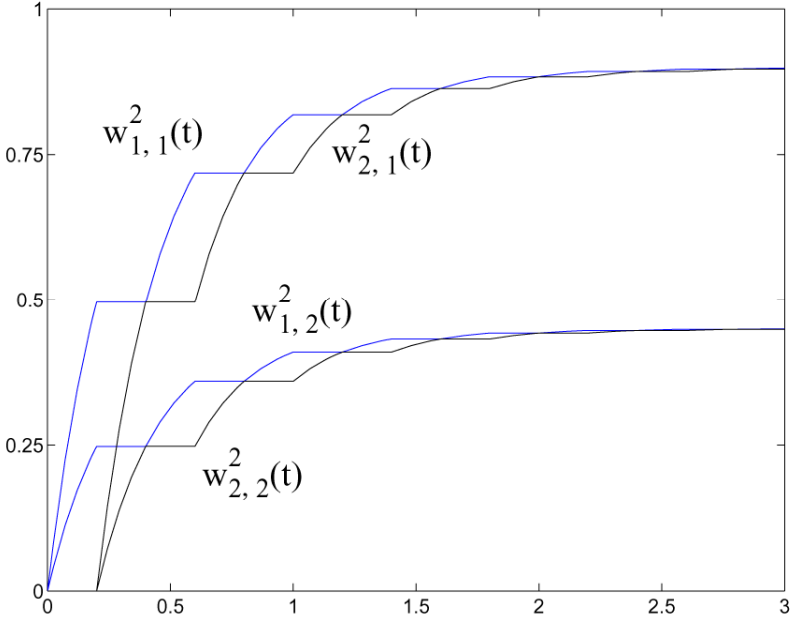
Two different input patterns are alternately presented to the network for periods of 0.2 seconds at a time.

For Pattern 1:

$$\mathbf{n}^1 = \begin{bmatrix} 0.9 \\ 0.45 \end{bmatrix} \qquad \mathbf{n}^2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

For Pattern 2:

$$\mathbf{n}^1 = \begin{bmatrix} 0.45 \\ 0.9 \end{bmatrix} \qquad \mathbf{n}^2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$




The first row of the weight matrix is updated when $n_1^2(t)$ is active, and the second row of the weight matrix is updated when $n_2^2(t)$ is active.

# شبکه‌ی گراسبرگ

## پاسخ وزن‌های وفقی

## RESPONSE OF ADAPTIVE WEIGHTS

فرض می‌کنیم دو الگوی ورودی متفاوت در دوره‌های زمانی 0.2 ثانیه‌ای به شبکه اعمال شوند.
فرض می‌کنیم لایه‌های ۱ و ۲ در مقایسه با وزن‌ها خیلی سریع‌تر همگرا می‌شوند و خروجی نرون‌ها در هر 0.2 ثانیه ثابت می‌شود.

Two different input patterns are alternately presented to the network for periods of 0.2 seconds at a time.

For Pattern 1:

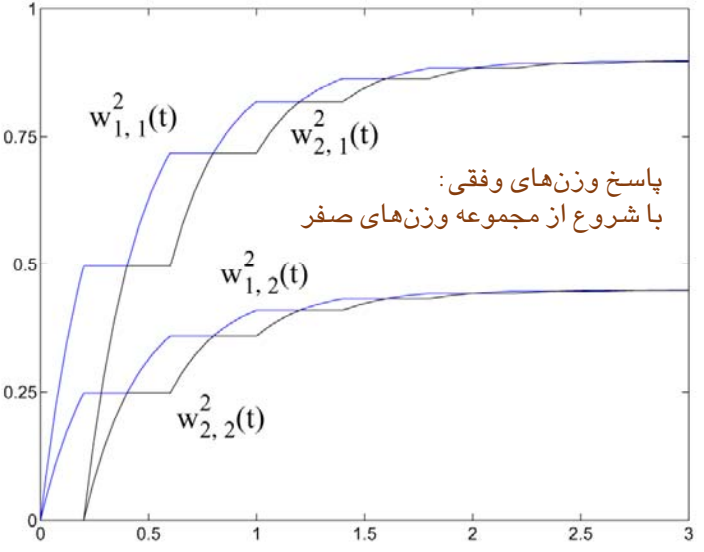
خروجی‌های لایه‌ی ۱ و ۲ برای دو الگوی ورودی

$$\mathbf{n}^1 = \begin{bmatrix} 0.9 \\ 0.45 \end{bmatrix} \quad \mathbf{n}^2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

For Pattern 2:

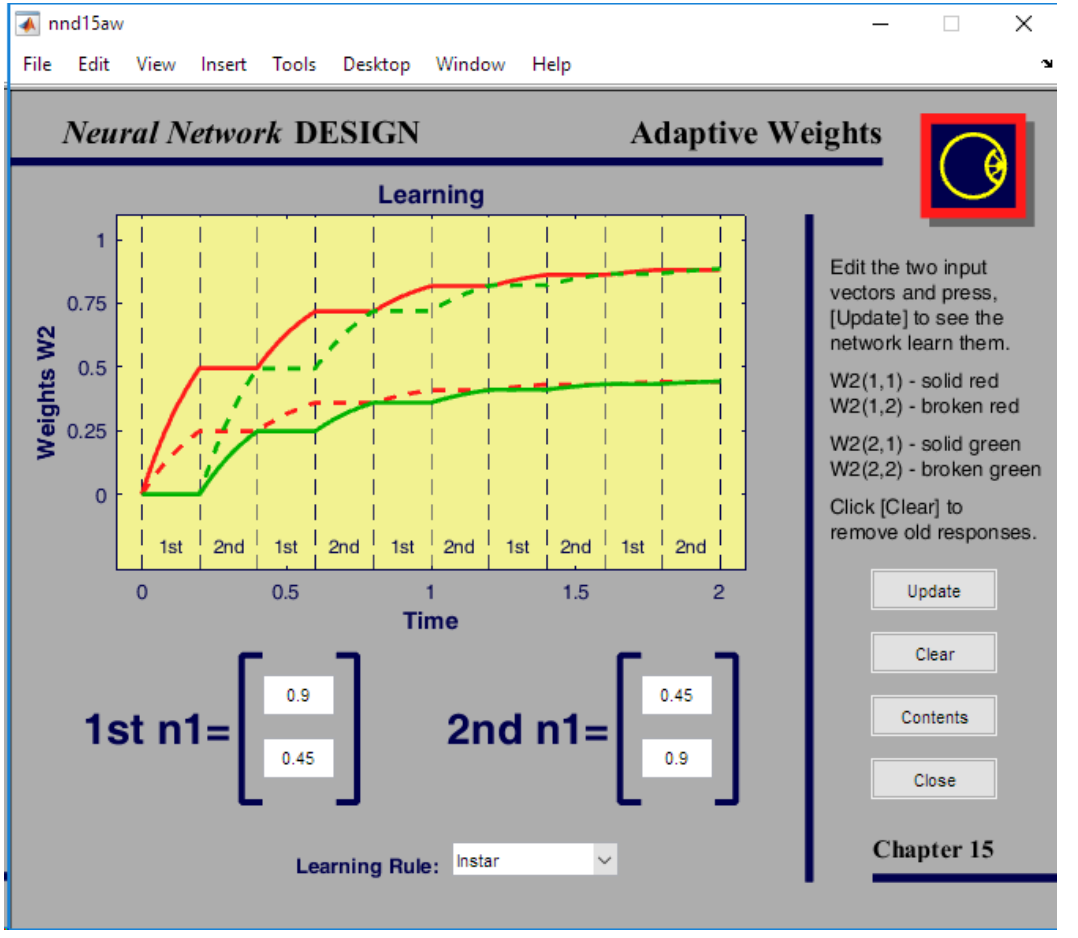
$$\mathbf{n}^1 = \begin{bmatrix} 0.45 \\ 0.9 \end{bmatrix} \quad \mathbf{n}^2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

الگوی اول توسط نرون اول از لایه‌ی ۲ و
الگوی دوم توسط نرون دوم از لایه‌ی ۲ کد شده است.




The first row of the weight matrix is updated when $n_1^2(t)$ is active, and the second row of the weight matrix is updated when $n_2^2(t)$ is active.

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




MATLAB

```
>> nnd15aw
```

شبکه‌ی گراسبرگ

# ۴

# رابطه با قانون کوهونن

# Relation to Kohonen Law

Grossberg Learning (Continuous-Time)

$$\frac{d[{}_i\mathbf{w}^2(t)]}{dt} = \alpha n_i^2(t)\{-[{}_i\mathbf{w}^2(t)] + \mathbf{n}^1(t)\}$$

Euler Approximation for the Derivative

$$\frac{d[{}_i\mathbf{w}^2(t)]}{dt} \approx \frac{{}_i\mathbf{w}^2(t+\Delta t) - {}_i\mathbf{w}^2(t)}{\Delta t}$$

Discrete-Time Approximation to Grossberg Learning

$${}_i\mathbf{w}^2(t+\Delta t) = {}_i\mathbf{w}^2(t) + \alpha(\Delta t) n_i^2(t)\{-{}_i\mathbf{w}^2(t) + \mathbf{n}^1(t)\}$$

# ارتباط قانون گراسبرگ با قانون کوهونن

## RELATION TO KOHONEN LAW

قاعده‌ی یادگیری گراسبرگ، نسخه‌ی پیوسته–زمان قاعده‌ی یادگیری Instar است.
به‌علاوه: شبکه‌ی گراسبرگ در ساده‌ترین فرم خود، نسخه‌ی پیوسته–زمان شبکه‌ی رقابتی کوهونن است.

### Grossberg Learning (Continuous-Time)

$$\frac{d[{}_i\mathbf{w}^2(t)]}{dt} = \alpha n_i^2(t)\{-[{}_i\mathbf{w}^2(t)] + \mathbf{n}^1(t)\}$$

### Euler Approximation for the Derivative

$$\frac{d[{}_i\mathbf{w}^2(t)]}{dt} \approx \frac{{}_i\mathbf{w}^2(t+\Delta t) - {}_i\mathbf{w}^2(t)}{\Delta t}$$

### Discrete-Time Approximation to Grossberg Learning

$${}_i\mathbf{w}^2(t+\Delta t) = {}_i\mathbf{w}^2(t) + \alpha(\Delta t) n_i^2(t)\{-{}_i\mathbf{w}^2(t) + \mathbf{n}^1(t)\}$$

# Relation to Kohonen Law

Rearrange Terms

$$ {}_i\mathbf{w}^2(t+\Delta t) = \{1-\alpha(\Delta t)n_i^2(t)\}\,{}_i\mathbf{w}^2(t) + \alpha(\Delta t)n_i^2(t)\{\mathbf{n}^1(t)\} $$

Assume Winner-Take-All Competition

$$ {}_{i^*}\mathbf{w}^2(t+\Delta t) = \{1-\alpha'\}\,{}_{i^*}\mathbf{w}^2(t) + \{\alpha\}'\mathbf{n}^1(t) \qquad \text{where} \qquad \alpha' = \alpha(\Delta t)n_{i^*}^2(t) $$

Compare to Kohonen Rule

$$ {}_{i^*}\mathbf{w}(q) = (1-\alpha)\,{}_{i^*}\mathbf{w}(q-1) + \alpha\mathbf{p}(q) $$

## ارتباط قانون گراسبرگ با قانون کوهونن

RELATION TO KOHONEN LAW

مرتب کردن مجدد جملات:

### Rearrange Terms

$$ {}_i\mathbf{w}^2(t+\Delta t) = \{1-\alpha(\Delta t)n_i^2(t)\}\,{}_i\mathbf{w}^2(t) + \alpha(\Delta t)n_i^2(t)\{\mathbf{n}^1(t)\} $$

برای ساده‌تر شدن تحلیل‌ها فرض می‌کنیم در لایه‌ی ۲ از تابع انتقال سریع‌تر از خطی استفاده شده باشد.

⇐ تنها یک نرون در لایه‌ی دوم دارای خروجی غیرصفر است: نرون $i^*$ ⇐ فقط سطر $i^*$-ام از ماتریس وزن‌ها به‌روزرسانی می‌شود.

### Assume Winner-Take-All Competition

$$ {}_{i^*}\mathbf{w}^2(t+\Delta t) = \{1-\alpha'\}\,{}_{i^*}\mathbf{w}^2(t) + \{\alpha\}'\mathbf{n}^1(t) \qquad \text{where} \qquad \alpha' = \alpha(\Delta t)n_{i^*}^2(t) $$

این معادله تقریباً با قاعده‌ی کوهونن برای شبکه‌های رقابتی یکسان می‌باشد.

بردار وزن مربوط به نرون برنده به سمت $\mathbf{n}^1$ حرکت خواهد کرد ($\mathbf{n}^1$ نسخه‌ی نرمال‌شده‌ی الگوی ورودی).

### Compare to Kohonen Rule

$$ {}_{i^*}\mathbf{w}(q) = (1-\alpha)\,{}_{i^*}\mathbf{w}(q-1) + \alpha\mathbf{p}(q) $$

# ارتباط قانون گراسبرگ با قانون کوهونن

مقایسه

RELATION TO KOHONEN LAW

سه فرق عمده بین شبکه‌ی گراسبرگ با شبکه‌ی رقابتی کوهونن:

- شبکه‌ی گراسبرگ یک شبکه‌ی پیوسته–زمان است
  (با مجموعه‌ای از معادلات دیفرانسیل غیرخطی توصیف می‌شود).

- لایه‌ی ۱ شبکه‌ی گراسبرگ به‌صورت خودکار بردارهای ورودی را نرمال می‌کند.

- لایه‌ی ۲ شبکه‌ی گراسبرگ یک رقابت نرم (soft competition) ایجاد می‌کند.
  در حالی که شبکه‌ی کوهونن رقابت «برنده همه را می‌خورد» دارد.
  رقابت نرم به تعداد بیشتری از نرون‌های لایه‌ی ۲ امکان یادگیری می‌دهد.

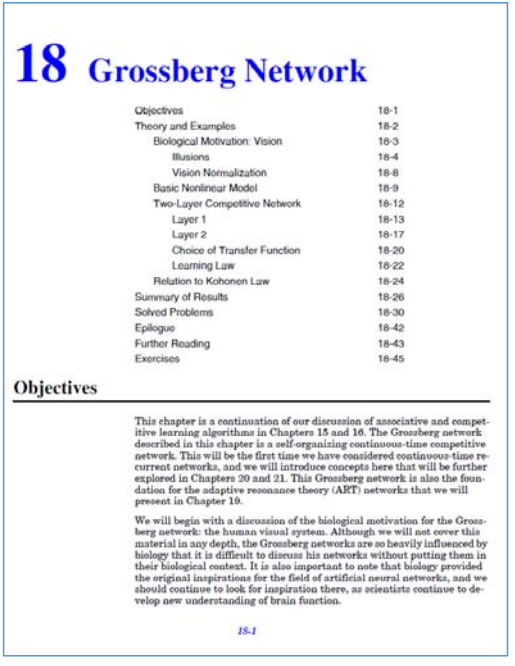
* شبکه‌ی گراسبرگ همانند یک شبکه‌ی نگاشت ویژگی عمل می‌کند.

شبکه‌ی گراسبرگ

# ۵

# منابع

# منبع اصلی



18 Grossberg Network

Objectives 18-1
Theory and Examples 18-2
Biological Motivation: Vision 18-3
Illusions 18-4
Vision Normalization 18-8
Basic Nonlinear Model 18-9
Two-Layer Competitive Network 18-12
Layer 1 18-13
Layer 2 18-17
Choice of Transfer Function 18-20
Learning Law 18-22
Relation to Kohonen Law 18-24
Summary of Results 18-26
Solved Problems 18-30
Epilogue 18-42
Further Reading 18-43
Exercises 18-45

Objectives

This chapter is a continuation of our discussion of associative and competitive learning algorithms in Chapters 15 and 16. The Grossberg network described in this chapter is a self-organizing continuous-time competitive network. This will be the first time we have considered continuous-time recurrent networks, and we will introduce concepts here that will be further explored in Chapters 20 and 21. This Grossberg network is also the foundation for the adaptive resonance theory (ART) networks that we will present in Chapter 19.

We will begin with a discussion of the biological motivation for the Grossberg network: the human visual system. Although we will not cover this material in any depth, the Grossberg networks are so heavily influenced by biology that it is difficult to discuss his networks without putting them in their biological context. It is also important to note that biology provided the original inspirations for the field of artificial neural networks, and we should continue to look for inspiration there, as scientists continue to develop new understanding of brain function.

18-1

Martin T. Hagan, Howard B. Demuth, Mark H. Beale, Orlando De Jesus,
**Neural Network Design**,
2nd Edition, Martin Hagan, 2014.
Chapter 18

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