

REDES NEURALES CAPACITIVAS.

Autor: Pr. Eric Fimbel

Departamento de Computación y Tecnología de la Información.

Universidad Simón Bolívar.

RESUMEN.

RESUMEN.

Hemos venido estudiando un tipo particular de redes: las redes de neuronas capacitivas.

Utilizan neuronas parecidas a condensadores que se cargan progresivamente y se descargan brutalmente.

Cada neurona tiene memoria de su pasado inmediato. Esa capacidad produce fenómenos temporales al nivel de la red completa: oscilaciones entre dos imágenes en caso de reconocimiento ambiguo, reconocimiento de secuencias temporales de imágenes, etc.

Por otra parte, utilizamos un sistema de crecimiento sináptico competitivo. Las sinapsis que llegan a una neurona tienen plasticidad, y compiten entre sí durante su crecimiento.

Cada neurona tiene una cantidad constante de un cierto recurso, repartido entre las sinapsis que llegan hacia ella. La fuerza de una sinapsis crece en forma no lineal con la cantidad de recurso que tiene (satura). Por lo tanto, a diferencia del modelo de Weber, la fuerza sináptica total no queda constante.

Hemos comprobado que este tipo de aprendizaje es robusto y bastante rápido.

Para experimentar, utilizamos redes sencillas. Son matrices de hasta 16 x 16 pixels, totalmente conectadas, sin nodos escondidos. Son más lentas que las redes de Hopfield, pero tienen una capacidad de memorización y reconocimiento mucho mayor.

I INTRODUCCION.

Hemos venido estudiando un tipo particular de redes: las redes de neuronas capacitivas.

Utilizan neuronas parecidas a condensadores que se cargan progresivamente y se descargan brutalmente.

Cada neurona tiene memoria de su pasado inmediato. Esa capacidad produce fenómenos temporales al nivel de la red completa: oscilaciones entre dos imágenes en caso de reconocimiento ambiguo, reconocimiento de secuencias temporales de imágenes, etc.

Por otra parte, utilizamos un sistema de crecimiento sináptico competitivo. Las sinapsis que llegan a una neurona tienen plasticidad, y compiten entre si durante su crecimiento.

Cada neurona tiene una cantidad constante de un cierto recurso, repartido entre las sinapsis que llegan hacia ella. La fuerza de una sinapsis crece en forma no lineal con la cantidad de recurso que tiene (satura). Por lo tanto, a diferencia del modelo de Weber, la fuerza sináptica total no queda constante.

Hemos comprobado que este tipo de aprendizaje es robusto y bastante rápido.

Para experimentar, utilizamos redes sencillas. Son matrices de hasta 16×16 pixels, totalmente conectadas, sin nodos escondidos. Son más lentas que las redes de Hopfield, pero tienen una capacidad de memorización y reconocimiento mucho mayor.

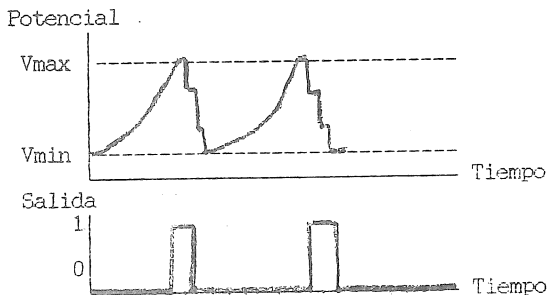
II. LAS NEURONAS CAPACITIVAS.

II.1 PRESENTACION.

Una neurona capacitiva tiene un POTENCIAL DE MEMBRANA V , que cambia con las SENALES excitatorias o inhibitorias que recibe.

Cuando el potencial pasa de un umbral V_{max} , la neurona emite un POTENCIAL DE ACCION, o sea una señal hacia las demás neuronas.

El potencial de acción rebaja el potencial de la neurona hacia un valor V_{min} , y el ciclo puede comenzar de nuevo.



La SALIDA de la neurona es un booleano que representa el potencial de acción. vale 1 cuando hay potencial de acción. 0 sino.

Se dice que la neurona está ACTIVA o PRENDIDA cuando su salida vale 1, APAGADA cuando vale 0.

El potencial de acción puede durar varios ciclos. Su duración es un parámetro de las neuronas.

Cuando no recibe señales, la neurona tiende a regresar a su POTENCIAL DE REPOSO, V_{min} . Es un proceso más lento que la emisión del potencial de acción. Pasa desapercibido cuando la neurona está activa.

II.2 MODELO BIOLOGICO SUBYACENTE.

Las neuronas animales tienen un potencial de membrana, o POLARIZACION. Cuando están inactivas, este potencial es negativo (≈ -60 MiliVoltios).

La neurona tiende a regresar a su potencial de reposo, gracias a un sistema de bombeo que mantiene las concentraciones internas de iones K^+ y Na^+ constantes (alta concentración de K^+ , baja concentración de Na^+).

Cuando recibe excitaciones por las sinapsis que le llegan, la neurona se DEPOLARIZA: su potencial de membrana aumenta (por ejemplo, sube desde -60 MiliVoltios hasta -40 MiliVoltios).

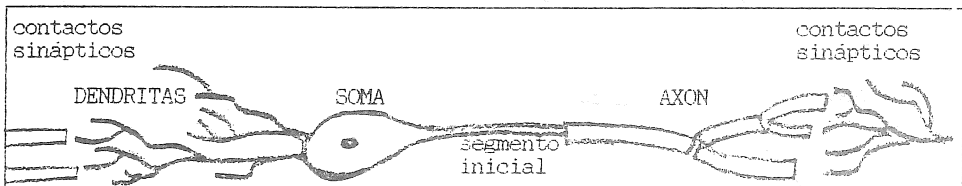
Al pasar el potencial de cierto umbral, se crea en la parte inicial del AXON (la "cola" de la neurona) un POTENCIAL DE ACCION SODICO.

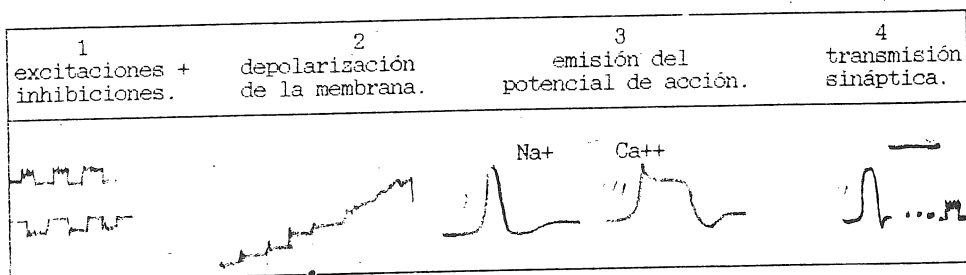
En el segmento inicial del AXON se produce una entrada brutal de iones Na^+ . Un poco más tarde, se produce una salida brutal de iones K^+ , y el potencial de la membrana vuelve a bajar (REPOLARIZACION).

El potencial de acción provocado por estos movimientos de iones se propaga en el axón, hacia las sinapsis. Allí, provoca el mando de una señal hacia las demás neuronas.

En los años setenta, se ha descubierto que las neuronas del sistema nervioso central podían emitir otro tipo de potencial de acción: el POTENCIAL DE ACCION CALCICO.

Este tipo de potencial, detectado anteriormente en ciertas fibras musculares, se debe a la entrada de iones Ca^{++} en vez de Na^+ . Presenta una duración mucho mayor que el potencial de acción sódico (de 10 a 30 milisegundos, versus 3 milisegundos).





II.3 LAS SINAPSIS.

Las sinapsis son los contactos que transmiten señales de una neurona hacia otra. tienen una FUERZA, que traduce la influencia de la señal sobre la neurona receptora. La fuerza sináptica puede variar durante los procesos de aprendizaje. eso explica la memorización de imágenes por parte de la red.

En nuestro modelo, existen sinapsis EXCITATORIAS e INHIBITORIAS. Desde la neurona i. hacia la neurona j, existen dos conexiones:

EXC(i,j), excitatoria ; INH(i,j), inhibitoria.

Las sinapsis son DISIMETRICAS. Si queremos mandar una señal de vuelta desde j hacia i. tenemos que utilizar dos sinapsis más:

EXC(j,i), excitatoria ; INH(j,i), inhibitoria.

También, existen AUTOCONEXIONES: son sinapsis excitatorias e inhibitorias desde una neurona hacia ella misma:

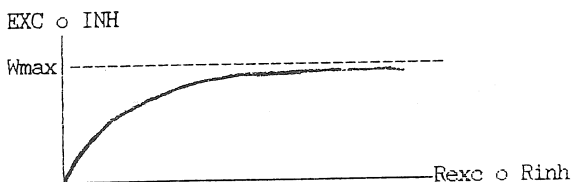
EXC(i,i), excitatoria ; INH(i,i), inhibitoria.

La fuerza sináptica resulta de un cierto RECURSO, presente en la neurona receptora. Cada sinapsis tiene una cierta cantidad de recurso, que determina su fuerza.

Notamos $R_{exc}(i,j)$ el recurso que tiene la sinapsis EXC(i,j).
Notamos $R_{inh}(i,j)$ el recurso que tiene la sinapsis INH(i,j).

La relación entre EXC() y $R_{exc}()$ por una parte, y entre INH() y $R_{inh}()$ por otra, es NO LINEAL. Para describirla, utilizamos una función sigmoide $W = \sigma(R)$:

$$W = 2 * W_{max} / (1 + e^{-(G * R)}) - W_{max}$$



II.4 MODELO BIOLOGICO SUBYACENTE.

Las sinapsis son contactos entre dos neuronas. Generalmente hacen contacto entre el axón de la neurona emisora y una ESPINA DENDRITAL de la neurona receptora.

La TRANSMISION SINAPTICA se hace como sigue:

Al recibir un potencial de acción, la neurona emisora libera un NEUROTRANSMISOR en el ESPACIO SINAPTICO.

La membrana de la neurona receptora contiene RECEPTORES, que capturan el neurotransmisor liberado.

Al capturar el neurotransmisor, cada receptor deja pasar una microcorriente, excitatoria o inhibitoria, según su dirección.

En el sistema nervioso central, existen autoconexiones (neuronas cuyos axones extienden una ramificación hacia su propio cuerpo).

También, existen sinapsis excitatorias y otras inhibitorias, según sea el neurotransmisor utilizado (Por ejemplo, el glutamato es excitador y el Gaba es inhibitor).

Cada neurona puede mandar sinapsis inhibitorias Y excitatorias a la vez. En efecto, se detectó que muchas neuronas utilizan varios neurotransmisores.

Por otra parte, las sinapsis de los organismos superiores son generalmente disimétricas. Existen sinapsis bidireccionales simétricas, o asimétricas (con fuerzas distintas). De todos modos, se modelan fácilmente mediante dos sinapsis, $EXC(i,j)$ y $EXC(j,i)$.

Por fin, la fuerza de una sinapsis depende basicamente de la cantidad de moléculas receptoras de neurotransmisor que contiene la membrana de la neurona receptora en el sitio de contacto.

En realidad, no sabemos si el recurso modela la concentración de receptores en la membrana o la concentración local de transductores del receptor (intervienen en su síntesis).

II.5 LAS ECUACIONES DE FUNCIONAMIENTO NEURAL: VERSION BASICA.

$V(j,t)$: Potencial de la neurona j al instante t .

$S(j,t)$: Salida de la neurona j al instante t . Es un booleano $\{0,1\}$.

$EXC(i,j)$: Fuerza de la sinapsis excitatoria desde i hacia j .

$INH(i,j)$: Fuerza de la sinapsis inhibitoria desde i hacia j .

$E(j,t)$: Señal total que recibe la neurona j al instante t .

Tenemos:

$$V(j,t+1) = V(j,t) * (1 - \beta) + E(j,t) - DV * S(j,t)$$

$$S(j,t+1) = (V(j,t+1) \geq V_{max}) * (1 - S(j,t)) + (V(j,t+1) \geq DV) * S(j,t)$$

$$E(j,t) = \sum_i (EXC(i,j) - INH(i,j)) * S(i,t)$$

Interpretación:

Las sinapsis excitatorias e inhibitorias se comportan como una sola sinapsis de fuerza $EXC(i,j) - INH(i,j)$. En cambio, durante el aprendizaje, ambas sinapsis crecen en forma independiente.

Cuando no recibe señales, el potencial tiende hacia 0, debido al coeficiente $(1 - \beta)$. 0 es el potencial de reposo de las neuronas.

Cada vez que la salida está en 1, el potencial baja de DV

La salida se coloca en 1 si:

- El potencial pasa de V_{max} o si
- La neurona ya estaba disparando ($S(j,t) = 1$) y su potencial permite seguir disparando ($V(j,t+1) \geq DV$).

Estas ecuaciones representan una aproximación inexacta de una neurona animal, pero permiten una implantación fácil sin alejarse demasiado de su modelo inicial.

II.6 DIFERENCIA CON LOS MODELOS CLASICOS DE NEURONAS.

Clasicamente, se usan neuronas sin memoria. Por ejemplo, la neurona bipolar clásica se dispara según la ecuación siguiente:

$$S(j,t+1) = (E(j,t+1) \geq 0)$$

En cambio, las neuronas capacitivas tienen memoria: el potencial V representa el resultado de las señales y disparos anteriores.

Observamos que el "grado de memoria" es ajustable. Si hacemos $\beta=1$, $DV=0$, $V_{max}=0$ en las ecuaciones, tenemos neuronas sin memoria:

$$S(j,t+1) = (E(j,t+1) \geq 0)$$

Por otra parte, las neuronas capacitivas tienen una característica importante:

No hay forma de determinar el estado interno a partir de la salida, debido a la discontinuidad de la función que determina la salida a partir del estado interno.

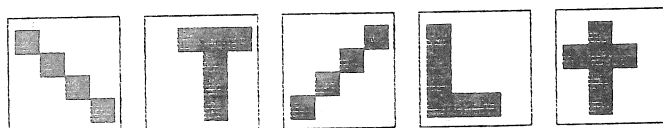
Nos parece que eso es la característica genérica que provoca el comportamiento particular de las redes capacitivas.

III. FUNCIONAMIENTO DE LA RED.

La red se compone de una matriz de hasta 16x16 puntos. Cada punto está conectado con todos los demás. El ejemplo presentado a continuación utiliza imágenes de 4x4 puntos.

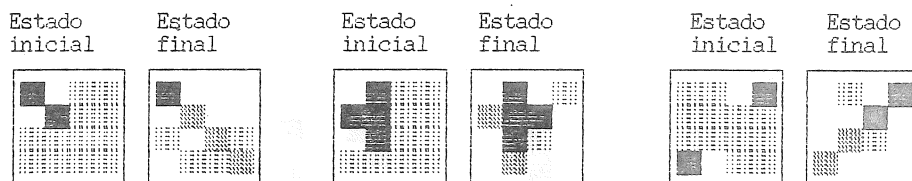
En un primer tiempo, presentamos a la red un conjunto de imágenes y la entrenamos para que las reconozca. Presentamos sucesivamente cada imagen a la red.

A cada presentación, se modifican las conexiones para grabar la configuración. Se repite el proceso 5 veces.



Luego, se da a la red una imagen parcial que deseamos que reconozca: la CONFIGURACION INICIAL. Los puntos tienen 3 estados posibles: Negro (prendido), Blanco (apagado) o Gris (indeterminado).

Dejamos correr la red, y vemos cuál es su ESTADO FINAL. En el estado final, cada punto tiene un tono de gris que describe su ACTIVIDAD (el promedio de sus salidas durante los 16 ciclos anteriores). Negro significa siempre prendido y blanco siempre apagado.



Observamos lo siguiente:

En el estado final, la imagen reconocida siempre aparece más oscura. Eso caracteriza un reconocimiento correcto.

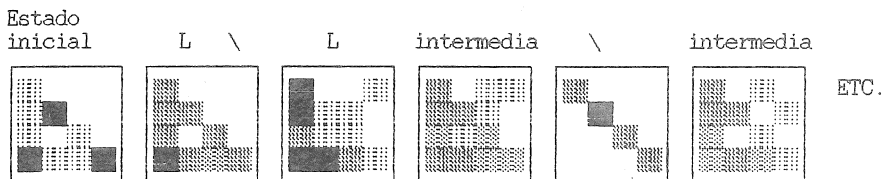
Para detectarlo formalmente, deberíamos utilizar una neurona fuera de la imagen, que se activaría solamente cuando el reconocimiento es correcto.

En los ejemplos, siempre hemos definido menos de 6 puntos (33% de la imagen). Eso basta para reconocer la imagen.

Además de la imagen reconocida, hay puntos parásitos que presentan una actividad (baja). Pertenecen casi siempre a imágenes que tienen puntos en común con la configuración inicial.

Eso se debe a que las neuronas presentan una ACTIVIDAD ALEATORIA: pueden activarse solas (ver más adelante).

En el ejemplo anterior, Las configuraciones iniciales corresponden a una sola imagen. Veamos ahora lo que pasa si la configuración inicial es ambigua, o sea corresponde a más de una imagen. El ejemplo siguiente es una configuración ambigua, parte "L" y parte "\".



Observamos lo siguiente:

La red está OSCILANDO entre dos configuraciones: la "L" y la "\".

Generalmente, se considera que las oscilaciones se deben a las conexiones negativas. No es cierto. Las pruebas hechas sin conexiones negativas muestran también oscilaciones.

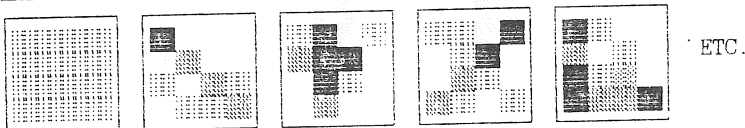
En ciertos casos, se estabiliza sobre una sola de las dos. Eso se debe a una mayor estabilidad de una de las dos configuraciones.

Las configuraciones más grandes son más estables, así como las que no solapan con ninguna otra configuración aprendida.

Casi nunca se llega a ver las dos al mismo tiempo en forma estable. Eso se debe a las conexiones negativas: al aparecer una configuración aprendida, tiende a apagar los nodos de las demás.

Por fin, veamos lo que pasa cuando se deja la red totalmente libre. La configuración inicial tiene todos sus nodos indeterminados:

Estado
inicial



Observamos lo siguiente:

Debido al RUIDO TERMICO (la actividad aleatoria de las neuronas), la red muestra en forma aleatoria las configuraciones aprendidas.

Al dejar funcionar la red cierto tiempo, se nota que la frecuencia de aparición de las imágenes no es homogénea.

Depende de la estabilidad propia de cada imagen. Las imágenes más grandes son más estables, así como las que no solapan con ninguna otra.

IV. ADICIONES AL MODELO DE NEURONA.

IV.1 EL RUIDO TERMICO Y LA ENERGIA DE UNA NEURONA.

El ruido caracteriza la actividad aleatoria de cada neurona. Se modela en la forma siguiente:

Cada neurona tiene excitaciones aleatorias de valor δV_{term} .

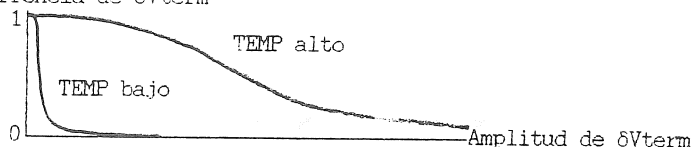
La probabilidad que ocurra la excitación δV_{term} baja a la medida que el AUMENTO DE ENERGIA DE LA NEURONA que provoca es más alto. La función de energía de la neurona se describe en anexo.

La probabilidad de que ocurra la excitación δV_{term} para la neurona j , al tiempo t , esta determinada mediante la ecuación:

$$\text{Proba}(\delta V_{term}) = e^{-(G * V(j,t) * \delta V_{term} / \text{TEMP})}$$

TEMP es la TEMPERATURA NEURAL de la red. Caracteriza la facilidad con la cual las neuronas cambian su potencial. Es un parametro EXOGENO. Veamos su influencia:

probabilidad de
ocurrencia de δV_{term}



Observamos lo siguiente:

Si $TEMP = 0$, no hay cambios aleatorios.

Si $TEMP$ es muy alto, todos los cambios pueden ocurrir con una probabilidad cerca de 1, la red no puede estabilizarse.

La probabilidad de tener cambios es baja cuando el potencial de la neurona es alto. El ruido térmico REGULA EL POTENCIAL (y por lo tanto la actividad) de las neuronas.

IV.2 PRESENTACION DE UNA CONFIGURACION A LA RED.

Los puntos que están prendidos en la configuración inicial pueden terminar apagados o semi apagados en la configuración final: la red tiene capacidad de corregir la configuración presentada. Eso se debe al método que utilizamos para presentar una configuración a la red:

Cuando un nodo está prendido en la configuración, recibe una EXCITACION SUPLEMENTARIA de valor δV_{exc} . Cuando está apagado, recibe una INHIBICION SUPLEMENTARIA de valor $-\delta V_{inh}$.

Si δV_{exc} y δV_{inh} son grandes ($\gg V_{max}$), entonces los nodos quedan en la configuración inicial. En cambio, si son pequeños, la red tiene mucha facilidad para "corregirlos".

Además, estas excitación e inhibición suplementarias no siempre se realizan. Por ejemplo, se hará solamente una vez cada 3 ciclos. En este caso, el PERIODO del refrescamiento es de 3.

Eso permite presentar a la red varias configuraciones en forma simultánea: presentamos cada una a su turno. Un periodo de 3 permite presentar 3 imágenes simultáneas.

En la práctica, un cambio de periodo no influye mucho sobre el reconocimiento. Si aumentamos el periodo, generalmente, se prenden y apagan nodos parásitos de vez en cuando.

IV.3 LAS ECUACIONES COMPLETAS DE FUNCIONAMIENTO DE LAS NEURONAS.

Describimos ahora las ecuaciones completas de funcionamiento, que integran ruido y presentación de imágenes.

Con las notaciones anteriores, tenemos ahora:

$$\begin{aligned} V(j,t+1) &= V(j,t) * (1 - \beta) + E'(j,t) - DV * S(j,t) \\ S(j,t+1) &= (V(j,t+1) \geq V_{max}) * (1 - S(j,t)) + (V(j,t+1) \leq -DV) * S(j,t) \\ E'(j,t) &= E(j,t) + \delta V_{term}(t) + \delta V_{exc}(j,t) - \delta V_{inh}(j,t) \\ E(j,t) &= \sum_i (EXC(i,j) - INH(i,j)) * S(i,t) \end{aligned}$$

$\delta V_{exc}(j,t) = \delta V_{exc}$ ssi j está prendido en la configuración inicial y hay exposición en el ciclo t . sino, es nulo.

$\delta V_{inh}(j,t) = \delta V_{inh}$ ssi j está apagado en la configuración inicial y hay exposición en el ciclo t . sino, es nulo.

V. EL APRENDIZAJE DE IMAGENES.

V.1 CONDICIONES DE CRECIMIENTO SINAPTICO.

Las fuerza de las sinapsis se modifica con la experiencia del sistema. Hay PLASTICIDAD SINAPTICA. Cuando se presenta una imagen nueva a la red, las sinapsis se modifican para grabarla.

Modelamos eso mediante una TEMPERATURA SINAPTICA. Lo más alta esta temperatura, lo más alta la probabilidad de que crezcan las sinapsis.

Cada categoria de sinapsis tiene su propia temperatura. Actualmente, las temperaturas son parámetros exógenos, manejados por el programa:

Texc es la temperatura de las sinapsis excitatorias.
Tinh es la temperatura de las sinapsis inhibitorias.

La probabilidad de que una sinapsis cualquiera $EXC(i,j)$ o $INH(i,j)$ se ponga a crecer depende de varios factores:

La probabilidad crece con la actividad de la neurona i (promedio de sus últimas salidas).

La probabilidad es más alta si la sinapsis estuvo creciendo al ciclo anterior (inercia de crecimiento).

Además, si se trata de una sinapsis excitatoria, $EXC(i,j)$:

La probabilidad crece con la temperatura sináptica Texc.

La probabilidad CRECE con el potencial $V(j)$.

La probabilidad baja con la cantidad de recurso $R_{exc}(i,j)$.

Si se trata de una sinapsis inhibitoria, $INH(i,j)$:

La probabilidad crece con la temperatura sináptica Tinh.

La probabilidad DECRECE con el potencial $V(j)$.

La probabilidad baja con la cantidad de recurso $R_{inh}(i,j)$.

Para escoger las sinapsis que crecen, utilizamos un método de Monte Carlo presentado en anexo.

V.2 EL MODELO BIOLOGICO SUBYACENTE.

Estamos modelando un tipo de crecimiento sináptico que no pasa de ser una hipótesis. Se basa en tres ideas:

Los receptores de un neurotransmisor están atraídos por el neurotransmisor o una sustancia liberada con el neurotransmisor.

Lo más activa está una sinapsis, lo más alta está la densidad promedio del neurotransmisor o la substancia en el espacio sináptico.

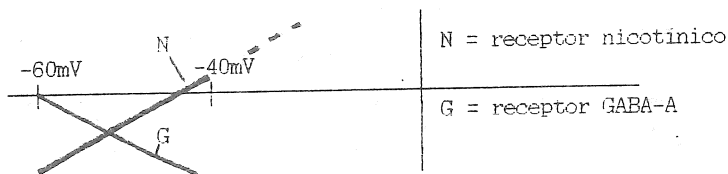
La probabilidad de atraer más receptores decrece con la densidad de receptores ya presentes.

Eso corresponde a una saturación, por cualquier razón que sea: repulsiones entre receptores, etc.

Los receptores migran más fácilmente en la membrana hacia las sinapsis más activas, si no están trabajando.

El receptor de un neurotransmisor inhibitor (por ejemplo, el GABA-A) trabaja poco si el potencial de membrana está bajo (la corriente que atraviesa el receptor es casi nula).

El receptor de un neurotransmisor excitator (por ejemplo, el receptor nicotínico de la acetilcolina) trabaja poco si el potencial de membrana está alto (la corriente que atraviesa el receptor es casi nula).



V.3 EL APRENDIZAJE COMPETITIVO.

La fuerza sináptica depende de la cantidad de recurso concentrado por la sinapsis. El recurso existe en cantidades limitadas en las neuronas: por lo tanto las sinapsis compiten para conseguirlo.

Cada categoría de sinapsis, positiva y negativa, utiliza un recurso distinto. Las sinapsis de una categoría compiten entre sí, y no interfieren con las sinapsis de la otra categoría.

En lo que sigue, consideramos las sinapsis excitatorias. El caso de las inhibitorias es idéntico.

Cada neurona tiene una cantidad fija de recurso: $R(j)$.

Cada sinapsis hacia j tiene una cantidad de recurso, $R(i,j,t)$, que depende del tiempo t .

El recurso que no pertenece a ninguna sinapsis está LIBRE. es $R_{lib}(j,t)$, que también depende del tiempo. Tenemos, a cada momento:

$$R(j) = R_{lib}(j,t) + \sum_i R(i,j,t)$$

En un principio, las sinapsis tienen poco recurso. $R(i,j,0) \approx 0$, y queda mucho recurso libre, $R_{lib}(j,0) \approx R(j)$. Si la sinapsis (i,j) crece, lo hace mediante la captura de un poco de recurso libre:

$$\begin{aligned} R(i,j,t+1) &= R(i,j,t) + \delta R \\ R_{lib}(j,t+1) &= R_{lib}(j,t) - \delta R \end{aligned}$$

Después de la etapa inicial de crecimiento, no queda recurso libre. $R_{lib}(i,j,t) \approx 0$. Cuando una sinapsis (i,j) crece, debe quitar recurso a otra sinapsis que llega a la misma neurona, (k,j) . HAY COMPETENCIA:

$$\begin{aligned} R(i,j,t+1) &= R(i,j,t) + \delta R \\ R(k,j,t+1) &= R(k,j,t) - \delta R \end{aligned}$$

La sinapsis (k,j) se escoge en forma aleatoria. Pero, en este sorteo, las sinapsis débiles tiene mayor probabilidad de ser víctimas que las sinapsis fuertes.

La implantación de este sorteo se hace mediante un algoritmo de Monte Carlo presentado en anexo.

V.4 DIFERENCIAS Y PARECIDOS CON LOS MODELOS CLASICOS DE APRENDIZAJE.

Los dos modelos más conocidos de aprendizaje son los de HEBB y WEBER. El modelo de Hebb dice que deben crecer las sinapsis entre nodos i y j conjuntamente prendidos en la configuración a aprender.

Esta estrategia minimiza la energía de la red para la imagen que está aprendiendo. Eso es cierto solamente para las redes sin autoconexiones, y con conexiones simétricas.

Nuestro criterio de crecimiento sináptico es una generalización del criterio de Hebb: la sinapsis excitatoria $EXC(i,j)$ crece con mayor probabilidad si i está activa y el potencial de j es alto.

Si i está activa, tiene una alta probabilidad de estar prendida (la actividad es el promedio de los últimos estados).

Por otra parte, Si el potencial $V(j,t)$ es alto, hay una fuerte probabilidad para que j se prenda pronto.

Las conexiones inhibitorias no están contempladas por el criterio de Hebb.

El modelo de WEBER dice que la sumatoria de todas las sinapsis que llegan a un nodo debe quedarse constante:

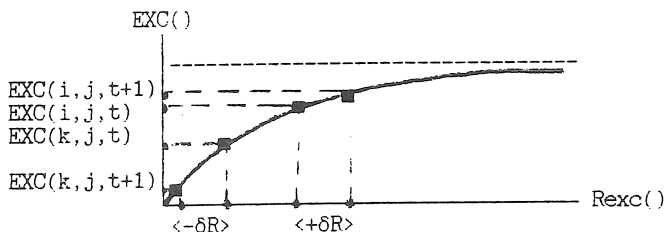
$$\sum_i W(i,j,t) = W(j) \text{ constante.}$$

Este criterio introduce la competencia entre sinapsis, pero directamente al nivel de la fuerza sináptica.

Por ejemplo, en caso de que $W(i,j)$ crezca:

$$\begin{aligned} W(i,j,t+1) &= W(i,j,t) + \delta W \text{ debe acompañarse de:} \\ W(k,j,t+1) &= W(k,j,t) - \delta W \end{aligned}$$

En cambio, la competencia para el recurso no mantiene constante la sumatoria de las fuerzas sinápticas. Por ejemplo, si la sinapsis (i,j) aumenta su fuerza, y quita recurso a la sinapsis (k,j) :



V.5 EL MODELO BIOLOGICO SUBYACENTE.

El RECURSO es la cantidad de moléculas receptoras del neurotransmisor en la membrana de la neurona receptora (puede ser el receptor mismo, o un transcriptor del receptor).

La unión sináptica que nos sirve de modelo para el aprendizaje competitivo es la unión nervio motor vs. célula muscular. Su desarrollo epigenético ha sido bastante estudiado.

En una primera etapa, llega un axón de motoneurona cerca de la fibra muscular. Establece una primera unión sináptica. Por debajo de esta unión, comienza a concentrarse el receptor (nicotínico) que hasta el momento estaba uniformemente diseminado en toda la membrana de la fibra muscular.

Llegan más y más axones de motoneuronas (hasta 30). Debido a un tropismo, van directamente hacia la primera sinapsis, y aquí establecen su sinapsis. En esta etapa, crecen sin competir.

Debido a una actividad motor automática (los fetos se mueven), las sinapsis trabajan. En esta etapa, compiten.

Al final de esta etapa, queda una sola: la más fuerte. Las demás están completamente eliminadas. La que queda no siempre es la primera: es la más fuerte.

Se supone, sin haberlo comprobado hasta el momento, que en el sistema nervioso central ocurre también el fenómeno de competencia sináptica. Esta competencia está por supuesto limitada a las sinapsis que utilizan un mismo neurotransmisor. Ahora, ¿en que medida influye la topografía compleja de una neurona sobre este fenómeno?

V.6 DETERMINACION DE LA CANTIDAD DE RECURSO INICIAL.

La cantidad de recurso inicial es un parámetro sumamente importante para el funcionamiento global de la red. Para la aplicación de reconocimiento de imágenes:

Si existe poco recurso, pocas sinapsis podrán crecer. Por lo tanto, el número y tamaño de las imágenes que se pueden aprender son muy limitados.

Si existe demasiado recurso, no hay competencia sináptica. Durante el reconocimiento, demasiados nodos tienden a prenderse, la red no discrimina bien entre las imágenes.

V.7 DETERMINACION DE LAS TEMPERATURAS SINAPTICAS.

Como hemos dicho anteriormente, en el protótipo actual, las temperaturas son parámetros exógenos. Son muy importantes: ya que definen los momentos durante los cuales la red aprende, cambiar completamente su comportamiento. Presentamos aquí dos heurísticas, que corresponden a dos versiones distintas del protótipo.

Aprendizaje separado de la corrida:

Los modos APRENDIZAJE y CORRIDA están separados. En el modo APRENDIZAJE, las temperaturas sinápticas $Texc$ y $Tinh$ están en alto, la red puede aprender. Durante la CORRIDA, están en 0.

Aprendizaje de asociaciones temporales (imagen A, luego imagen B):

Se aprende A y B por separado como visto en el caso anterior.

Luego, se presenta A y se deja estabilizar la red.

Luego, se presenta B, y se pone en alto la temperatura de las sinapsis excitatorias, $Texc$, durante pocos ciclos.

Crecen las conexiones excitatorias entre los nodos de A y los de B: la imagen A tiende a prender la imagen B.

Veamos un ejemplo de reconocimiento de asociación: '\', luego '/'.

Se configura una parte de la imagen inicial (50 %).
Después de C ciclos, se libera completamente la imagen.
La segunda imagen aparece superpuesta con la primera.
La segunda elimina la primera y queda sola.



VI. CONCLUSIONES PROVISIONALES.

Las redes capacitivas parecen tener una alta capacidad de almacenamiento (hasta 8 patrones en una red de 4 x 4 puntos), así como un buen nivel de reconocimiento (a veces, basta prender un solo nodo para reconocer una imagen completa).

Los algoritmos de aprendizaje son bastante robustos: el resultado es independiente del orden de presentación de las imágenes, el número de pasadas necesarias es bajo (nunca ha necesitado más de 6 pasadas).

Por fin, el aprendizaje de asociaciones temporales (A luego B) es uno de nuestros mayores logros: sin nodos ni conexiones suplementarios, con un algoritmo sencillo (el mismo algoritmo de aprendizaje sirve), la red reconoce asociaciones temporales en forma bastante buena.

En cambio, si la convergencia es rápida, los tiempos de cálculo son mayores a los de las redes de hopfield. Hace falta optimizar los algoritmos y estructuras de datos utilizados

Además, por el momento, la red no tiene sistema de salida: el ojo humano ve perfectamente si la red ha reconocido una imagen o no, pero todavía no hemos implantado un sistema automático para hacerlo.

AGRADECIMIENTOS.

A J.F. ARIAS, C. CONTRERAS, G. GONZALEZ, G. HERNANDEZ DE VIVAS, P.L. ROMERO, V. YRIARTE, y a todos los que han participado a los desarrollos de programas, los experimentos y el establecimiento de los modelos de redes capacitivas.

BIBLIOGRAFIA.

- ABU MUSTAFA Y.S.
"Connectivity versus entropy", Neural information processing, Denver, PP.1.8, 1988.
- ACKLEY D.H., HINTON G. E., SEJNOWSKY T.J.
"A learning algorithm for Boltzmann machines" Cognitive science V.9, pp. 147-160, 1985.
- BORG- GRAHAM L.J.
"Simulations suggest information processing roles for the divers currents in hippocampal neurons", Neural information processing, Denver, PP.1.8, 1988.
- GUEVARA M., GLASS L., MACKEY M. SCHRIER A.
"Chaos in neurobiology", IEEE PP.790 798, 1983.
- HAMMOND C., TRITSCH D.
"Neurobiologie cellulaire, Doin Editeur, 1990
- HUBEL D.H., WIESEL T.N.
"Sequence regularity and geometry of orientation columns in the monkey striata cortex" Journal of comparative Neurology 158. PP. 267-274, 1974.
- KLARSEFELD A.
"La compartimentation des fibres musculaires", La recherche v. 228, pp. 116-119, 1991.
- KOSKO B.
"Bidirectional associative memories" IEEE Ttransactions on systems, vol 18 PP.49-60, 1988.
- LANG K.J., HINTON G.E.
"The development of the time delay Neural Network Architecture for speech recognition" Tech.Rep. CMUCS, pp. 152, Carnegie Mellon University, 1988.
- LAUFER R. CHANGEUX J.P.
"Activity dependant regulation of gene expression in muscles and neuronal cells" Mol. Neurobiol, t3, pp. 1-53, 1989.
- LAUTRUP B.
"Monte Carlo Methods in theoretical high-energy physics", ACM, V. 28, PP. 358 362., 1985.
- NAKAMURA K., ICHIKAWA A.
"Timing information in transient behavior of neuropopulations", IEEE transactions on systems, v.19, pp 1188-1192, 1989.
- VON DER MALSBURG C.
"Self organization of sensitive cells in the striata cortex" Kibernetik vol 14. PP 85-100, 1973.

VII. ANEXO.

VII.1 ECUACIONES DEL RUIDO TERMICO.

La variación de energía debida a una excitación se aproxima mediante el siguiente modelo:

Consideramos la neurona como un condensador. En este caso, tenemos la siguiente relación entre su potencial $V(j,t)$ y su carga eléctrica $Q(j,t)$:

$$V(j,t) = Q(j,t) / \text{Capacidad}(j)$$

En primera aproximación consideramos una excitación como una entrada de carga δQ_{term} que provocaría una variación de potencial δV_{term} si la neurona estuviese aislada.

$$\delta V_{\text{term}} = \delta Q_{\text{term}} / \text{Capacidad}(j)$$

La energía a gastar para realizar esta variación de carga es:

$$\delta \text{Energia} = V(j,t) * \delta Q_{\text{term}}$$

$$\delta \text{Energia} (\delta V_{\text{term}}) = V(j,t) * \delta V_{\text{term}} * \text{Cste.}$$

La probabilidad de aceptar la excitación δV_{term} es dada por la distribución de Boltzmann:

$$\text{Proba}(\delta V_{\text{term}}) = e^{**} (- G * \delta \text{Energia} (\delta V_{\text{term}}) / \text{Temp})$$

VII.2 EL METODO DE MONTE CARLO.

Se utiliza en varios procesos: el ruido térmico, la competencia sináptica, etc. Permite tomar una decisión con una probabilidad dada. El problema no es simple: una probabilidad se establece generalmente a partir de un gran número de experimentos. El método general es el siguiente:

Calcular la probabilidad PRB de tomar una decisión.

Escoger un número UMB, al azar en $[0..1]$, con una ley uniforme.

Si $\text{PRB} \geq \text{UMB}$ se acepta la decisión.

En el caso contrario, se la rechaza.

VII.3 EL RUIDO TERMICO.

Excitación aleatoria de la neurona j , al tiempo t :

Escogemos al azar un δV_{term} .

Calculamos $P = \text{Probabilidad} (\delta V_{\text{term}})$.

Escogemos un número UMB, al azar en $[0..1]$, con una distribución uniforme.

Si $P \geq \text{UMB}$, se realiza la excitación.

En caso contrario, no pasa nada.

VII.4 CRECIMIENTO SINAPTICO.

Se escoge las sinapsis $EXC(i,j)$ que van a crecer en la forma siguiente:

Se chequean todos los nodos i tales que $actividad(i) \geq 0.7 * 16$ (16 es el máximo de las actividades).

Escogemos un nodo j al azar. Si $V(j,t) < V_{max} * 0.7$, no hace nada. V_{max} es el valor máximo del potencial de membrana.

Calculamos $PRB = actividad(i) * V(j,t) * (R_{max} - R_{exc}(i,j)) * Texc$
 R_{max} es el valor máximo del recurso de una sinapsis.

Si $EXC(i,j)$ estaba creciendo al tiempo $t-1$, $PRB = PRB * 2$.

Normalizamos PRB en $[0..1]$.

Escogemos un número UMB , al azar en $[0..1]$, con una distribución uniforme.

Si $PRB > UMB$, se aumenta la sinapsis $EXC(i,j)$.
Sino, no pasa nada.

Las sinapsis inhibitorias se manejan en la misma forma, salvo que:

Se descarta el nodo j si $V(j) > 0.3 * V_{max}$.

$PRB = actividad(i) * (V_{max} - V(j,t)) * (R_{max} - R_{exc}(i,j)) * Tinh$

VII.5 COMPETENCIA SINAPTICA.

Dado el nodo j , ya hemos escogido una sinapsis (i,j) que va a aumentar su fuerza. debemos escoger una sinapsis (k,j) que sea la victima, o sea tal que le quitamos su recurso.

Escogemos al azar un nodo k .

Normalizamos $EXC(k,j)$ en $[0..1]$. llamemos EXC el valor normalizado.

Escogemos un número UMB , al azar en $[0..1]$, con una distribución uniforme.

Si $EXC < UMB$, se realiza la excitación.

En caso contrario, se itera el proceso de búsqueda de victima.